A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles

Predicting the Direction of Selective Recall: Its Relation to Ego Strength and N Achievement 149
Thelma G. Alper

The Relationship Between Advanced Age and Mental Abilities 166
Charles R. Strother, K. Warner Schaie, and Paul Horst

Judgments of Consensus During Group Discussion 171
Phil Stone and Joe Kamiya

Cognitive Processes and Interpersonal Predictions 176
Howard Leventhal

Conforming Behavior Under Two Conditions of Conflict in the Cognitive Field 181
Frans Samelson

The Repression Response to Implied Failure as a Function of the Hysteria-Psychasthenia Index 188
Charles B. Truax

Intra-Individual Consistency in Conceptualization 194
Louis J. Moran, Lawrence S. McGaughran, and Don Leventhal

Continued on inside front cover

Volume 55 · SEPTEMBER 1957 Number 2

CONTENTS—*Continued*

Distinguishing Characteristics of Collaborators and Resisters Among American Prisoners of War 197
Edgar H. Schein, Winfred F. Hill, Harold L. Williams, and Ardie Lubin

The Relationship of Age and Privilege Status to Reaction Time Indices of Schizophrenic Motivation 202
Gerald Rosenbaum, James L. Grisell, and William R. Mackavey

The Effects of Sex-Role Identification Upon Problem-Solving Skill 208
G. A. Milton

The Effects of Clear and Unclear Role Expectations on Group Productivity and Defensiveness 213
Ewart E. Smith

The Relationship Between Overt and Fantasy Aggression as a Function of Maternal Response to Aggression 218
Gerald S. Lesser

Affiliation Motivation and Productivity in Small Groups 222
Richard deCharms

Cognitive Set and Motivational Factors in the Perception of Neutral and Threat-Related Stimuli 227
Ruth C. Wylie

Maternal Child-Rearing Practices and Authoritarian Ideology 232
I. Hart

Catharsis and the Relief of Hostility 238
Philip Worchel

Assimilation and Contrast Effects in Reactions to Communication and Attitude Change 244
Carl I. Hovland, O. J. Harvey, and Muzafer Sherif

Compatriot and Foreigner: A Study of Impression Formation in Three Countries 253
Jerome S. Bruner, and Howard V. Perlmutter

CASE REPORT

The Treatment of Two Phobic Patients on the Basis of Learning Principles .. 261
Victor Meyer

CRITIQUE AND NOTES

Perception of Disturbing and Neutral Words Through the Autokinetic Word Technique 267
Sarnoff A. Mednick, Alan Harwood, and Jack Wertheim

Learning Theory and "Opposite Speech" 268
Arthur W. Staats

The Influence of Ethnic Attitudes on Reasoning About Ethnic Groups 270
Norman M. Prentice

ERRATA 273

Volume 55 SEPTEMBER, 1957 Number 2

The Journal of Abnormal and Social Psychology

Editor: M. BREWSTER SMITH, *New York University*

Associate Editor: E. J. SHOBEN, JR., *Teachers College, Columbia University*

CONSULTING EDITORS

GORDON W. ALLPORT
Harvard University

URIE BRONFENBRENNER
Cornell University

DONALD T. CAMPBELL
Northwestern University

LAUNOR F. CARTER
The RAND Corporation

RICHARD CHRISTIE
Columbia University

MORTON DEUTSCH
Bell Telephone Laboratories

I. E. FARBER
State University of Iowa

LEON FESTINGER
Stanford University

DONALD W. FISKE
University of Chicago

ROBERT R. HOLT
New York University

HERBERT H. HYMAN
Columbia University

IRVING L. JANIS
Yale University

HAROLD E. JONES
University of California

HAROLD H. KELLEY
University of Minnesota

LEONARD S. KOGAN
Community Service Society of New York

GARDNER LINDZEY
University of Minnesota

DAVID C. MCCLELLAND
Harvard University

QUINN MCNEMAR
Stanford University

LEO POSTMAN
University of California

ELIOT H. RODNICK
Duke University

NEVITT SANFORD
University of California

SEYMOUR SARASON
Yale University

RICHARD L. SOLOMON
Harvard University

RENATO TAGIURI
Harvard University

Managing Editor: ARTHUR C. HOFFMAN

Assistant Managing Editor: HELEN ORR

Editorial Staff: BARBARA CUMMINGS, FRANCES H. CLARK, SADIE J. DOYLE, SARAH WOMACK

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued bimonthly in January, March, May, July, September, and November, two volumes per year. The two volumes average 784 pages. The subscription price per year is $16.00; foreign, $16.50. Single issues are $3.00.

Original manuscripts should be sent to the Editor:

DR. M. BREWSTER SMITH
Graduate Department of Psychology
New York University
Washington Square
New York 3, N. Y.

Fifty free offprints are supplied to contributors. Authors of prior publication articles receive no gratis offprints.

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Printed and mailed by:

WAVERLY PRESS, INC.
Baltimore 2, Maryland

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Mt. Royal and Guilford Aves., Baltimore 2, Md.
and 1333 Sixteenth Street N.W., Washington 6, D.C.

Entered as second-class matter February 15, 1954 at the Post Office at Baltimore, Maryland under the Act of February 28, 1925. Acceptance for mailing at special rate of postage provided for in Section 34.40 (e), P. L. and R. Additional entry at Washington, D.C.

Copyright 1957 by the American Psychological Association, Inc.

PREDICTING THE DIRECTION OF SELECTIVE RECALL: ITS RELATION TO EGO STRENGTH AND *N* ACHIEVEMENT[1]

THELMA G. ALPER

Wellesley College

WHEN the research findings in a given area continue over the course of many years to be as diverse as those on the recall of interrupted vs. completed tasks, the challenge to find out why can be strong. This has been the motivation behind the series of four papers by the writer (1, 2, 3, 4), the first of which appeared in 1946. At the time the research for the first two papers was undertaken, several studies had appeared. All of them had been designed to test further the conditions under which the Zeigarnik (29) finding, that interrupted tasks are more likely to be recalled than completed ones, would hold. Many of these studies, as previous reviews of the literature have shown (e.g., 1, 3, 13, 26), seemed not to support Zeigarnik since they did not yield statistically significant recall differences in the expected direction under conditions where completion and incompletion of tasks was experimentally controlled. One experiment, in fact, (24) indicated a reversal in the recall direction, i.e., a greater recall of completed than of interrupted tasks. The subjects (*S*s) who served in that experiment, a group of crippled children, had been told "that they were to be given a test to determine how well they could do puzzles and that a prize would be given the one who did best" (24, p. 249). Not all of the *S*s behaved in the same way. Some showed the reversal, some did not, others recalled an equal number of completed and interrupted tasks. The reversal pattern, however, was characteristic mainly of *S*s rated by their teachers to be high on "pride." Though the ratings were admittedly not highly reliable, this was one of the first studies to support Zeigarnik's suggestion that personality factors may play a part in selective recall (29, p. 303). Zeigarnik's finding here with adults was that "ambitious" subjects, identified by remarks such as "What is the matter with me—!" or "I wouldn't have believed I could be so stupid," recalled even *more* interrupted than completed tasks (29, p. 307).

The importance of another variable is also suggested by the Rosenzweig and Mason study; namely, the orientation of the experimental instructions. Working for a *prize* which you may get by doing well on a *test* is probably dynamically not the same as working on some tasks because someone asks you to do so.[2] Zeigarnik's instructions had been "I shall give you a series of tasks which you are to complete as rapidly and correctly as possible" (29, p. 300). In a later experiment, Rosenzweig (23) tested the orientation variable. He asked one group of college males, the "informal" group, to help the experimenter (*E*) classify some jigsaw puzzles in preparation for their use in a later experiment. Other *S*s, the "formal" group, were told that the puzzles were part of an intelligence test. The selective recall results of the two groups were found to differ: the informal group recalled significantly more interrupted tasks, the formal group significantly more completed tasks. In Zeigarnik's study, no attempt had been made to maintain a uniform experimental atmosphere. On the contrary, Zeigarnik tells us that in keeping with differences in the *S*s' reactions to the situation, wanting "to please the *E*," or "to excel as if in competition with others," or "interested in the task for its own sake" (29, p. 303), the *E* "did not preserve a fixed mien and method with all subjects. Those of the first type were allowed to see the experimenter's pleasure when the task was well done. Work by the second group was inspected with the air of an examiner, while the third group was allowed to work unmolested, the experimenter in this case remaining passive" (29, p. 303).

In the meantime, Alper (1), working also within Zeigarnik's tension theory of selective recall, and taking up Zeigarnik's (29, footnote 1, p. 312) suggestion that the recall of a given

[1] The writer acknowledges with thanks the invaluable help of Henry A. Murray, Robert W. White, and and Eugenia Hanfmann.

[2] See Alper (3), for a further discussion of this point with special reference to the sample used in the Rosenzweig and Mason (24) study.

S might differ in different experimental settings, proposed that a more sensitive way to test for the conditions that influence the direction of recall might be to measure the recall of the *same Ss* under two different experimental settings: task orientation in one session, ego orientation in the other. To create a task-oriented atmosphere, the instructions emphasized, as Rosenzweig (23) had done, that *E* was testing the materials, not the *S*. For the ego-oriented atmosphere, where the intention was to threaten self-esteem, *Ss* were told that this was a standardized set of tasks and that they constituted a brief intelligence test devised for the army to use in selecting candidates for officers' training school. The *Ss* were all draft age college males. As predicted, *Ss* differed in their reactions to both instructions, and there were no significant recall differences under either orientation. When the recall results were correlated with personality ratings derived from an intensive clinical study of these *Ss* within the framework of Murray's (21) need theory of personality, however, two major patterns of recall emerged: (*a*) the Strong Ego pattern, consisting of the greater recall of interrupted tasks under task-oriented conditions, and of completed tasks under ego orientation; and (*b*) the Weak Ego pattern, consisting of a greater recall of completed tasks under task-oriented conditions, and of interrupted tasks under ego orientation.

One other variable had been controlled in the above study—the nature of the tasks used. Alper and Black (4) have shown that most of the tasks used by Zeigarnik (29), Lewis (15), and others, would be unacceptable to college students as IQ tests. Or, to put it concretely, failure to complete within a given time limit tasks such as bead-stringing, winding thread on a spool, circling vowels, etc., would not be regarded by an *S* as a reflection on his intelligence, though failure to complete an anagram or a scrambled phrases task would. The use of tasks not commensurate with the atmosphere *E* intends by his instructions may account for some of the previously equivocal findings in this field, as the writer has pointed out elsewhere (1, 4). In the experiments referred to above, the writer had used a scrambled phrases task.

To test the validity of the Strong and Weak Ego recall patterns, the writer designed another experiment. Now two groups of *Ss*, a Strong Ego and a Weak Ego group, respectively, were selected in advance by means of specially constructed instruments. From the direction of their total recall patterns over two sessions, task-oriented first, and later an ego-oriented one, it was possible to predict correctly the group membership of these *Ss* as Strong or as Weak Egos significantly beyond chance. This experiment has been cited briefly elsewhere (2), but has not been reported in full.

In 1952, reacting to Sears' (26) recommendation that the interruption technique be abandoned as unsatisfactory since so little agreement had been reached by the various *Es*, the writer (3) tried to show: (*a*) that both orientation and personality were important in selective recall; and (*b*) that when both were considered, the seemingly diverse results of Zeigarnik (29), Rosenzweig and Mason (24), Rosenzweig (23), Glixman (13), Sanford (25), and others could be fitted into a common theoretical framework. The writer had intended to let the matter rest here. Recently, however, studies by Atkinson (5), and Atkinson and Raphelson (6) have renewed her interest since the data reported by these authors put a different interpretation on the results of earlier studies. They report that high *n* Achievers recall more interrupted than completed tasks under Achievement Orientation while low *n* Achievers show a reversed recall pattern. While these are the same recall patterns the writer has discussed, the personality relationships involved are not the same and the interpretations given them are also not the same. In fact, on the surface, at least, the *n* Achievement and Ego Strength frameworks would seem to have yielded dynamically opposite results.

Since *n* Achievement had not been measured in the writer's earlier studies, to test the relationship between *n* Achievement and Ego Strength requires a new experiment. This experiment has now been done and is reported in Part II of the present paper. In Part I, the writer has gone back to the prediction study mentioned above and is reporting it now in full since it confirms the validity of the Strong and Weak Ego recall patterns statistically revealed by the syndrome analysis used in the earlier study. It provides us, therefore, with a firmer base for testing the

relationship between *n* Achievement and Ego Strength and for applying the results to other selective recall studies.

PART I

PREDICTING THE DIRECTION OF RECALL OF STRONG AND WEAK EGO SUBJECTS

In the writer's earlier papers on the recall of interrupted vs. completed tasks (1, 2), the major premise had been that the direction of the recall ratio is a function of the personality structure of the individual recaller. It had been hypothesized "that the direction of selective recall of a given subject will differ in a non-self-esteem involving laboratory situation and in a laboratory situation where self-esteem is objectively threatened in a manner which is consistent with the self-esteem needs of that subject" (2, pp. 104–105). In designing the experiment to test this hypothesis, therefore, both instructional set and personality structure were taken into account.

Ten male undergraduates had served as *Ss*. Each was tested first in a task-oriented and later in an ego-oriented setting. The intention had been to threaten self-esteem objectively only in the latter setting. Since personality ratings, based on intensive clinical studies of these *Ss* within the framework of Murray's (21) need theory of personality were available, selective recall could be correlated with personality structure. Horn's (14) syndrome analysis technique was used for this purpose.

The analysis yielded two major selective recall patterns. One pattern, made up of the recall of more interrupted than completed tasks under Task Orientation and of more completed than interrupted tasks under Ego Orientation, was labeled the "Strong Ego" pattern since the syndrome analysis showed it to be characteristic only of *Ss* who had been rated high on *n* Recognition,[2] *n* Dominance,[4] Ego Strength,[5] and Conative Conjunctivity,[6] and low on Dejection Pessimism[7] and Ego-Ideal Intragression.[8] The opposite recall pattern, consisting of the recall of more completed than interrupted tasks under Ego Orientation, was labeled "Weak Ego" since it was characteristic only of *Ss* who had been rated high on Dejection-Pessimism and Ego-Ideal Intragression, and low on *n* Recognition, *n* Defendance,[9] *n* Counteractive Achievement,[10] and Narcism.[11]

In order to test whether these statistically derived patterns could be experimentally reproduced, the prediction study described here was designed.

Method

Two groups of nine *Ss* each, a Strong Ego and a Weak Ego group respectively, were selected in advance of the experiment by means of specially constructed clinical tools. All *Ss* were tested under Task Orientation first and later under Ego Orientation.

It was predicted that the recall patterns of these *Ss* would be consistent with the findings of the earlier experiment since, apart from the fact that personality structure now served as an independent variable, the design of the prediction study followed closely the plan of the previous experiment.

Selection of subjects. Two procedures were used to select *Ss*. First, a self-rating scale, *The Psychological Insight Test*, was constructed to measure the personality variables in the Strong and Weak Ego syndromes. Second, *Ss* whose self-ratings identified them as Strong or as Weak Egos were interviewed. Only *Ss* whose

---

existing or acquirable. This manifests itself chiefly as a successful *n* Achievement, giving proof of the power to persist; an unconquerable will.

[6] Conative Conjunctivity: to organize one's efforts; to make plans and to follow them; to force drives into an efficient pattern; to live an ordered life. This variable gives evidence of the 'coordinating powers of the ego.'

[7] Dejection Pessimism: to worry about one's shortcomings; to be pessimistic about life; to recover slowly from disappointment.

[8] Ego-Ideal Intragression: to feel humiliated and ashamed after failure; to be burdened by feelings of inferiority. The ego is belittled and criticized for stupidity after failure.

[9] *n* Defendance: to defend oneself against blame, belittlement, assault; to justify one's actions; to offer extenuations, explanations and excuses for oneself.

[10] *n* Counteractive Achievement: to set difficult goals for oneself and attempt to reach them. To stick at a task until satisfied with the results; to overcome obstacles; to restrive after failure; to redouble one's efforts in the face of frustration.

[11] Narcism: to perceive the world from a personal or subjective view-point; to be disdainful of others; to be dominated by ruthless self-seeking; to purport to be indifferent to one's failures and to one's critics; to be self-satisfied; to be eccentric.

[2] *n* Recognition: to seek to achieve something in order to win acclaim, renown, fame; to seek prestige and glory; to put oneself forward; to boast of one's accomplishments.

[4] *n* Dominance: to seek to control the sentiments and behavior of others by suggestion, seduction, persuasion or command; to lead others; to organize and direct groups of people; to get others to cooperate.

[5] Ego Strength: to know what one wants to do and has the capacity to do, and to do it. It is a matter of *willing* one's desires within the limits of one's abilities,

TABLE 1
THE PSYCHOLOGICAL INSIGHT TEST

A Sample of the First Twelve Items

1. I enjoy acting before a large audience. (*n* Recognition)
2. I prefer to rely on my own judgments rather than depend on other men's opinions. (Narcism)
3. I look upon failure as a challenge to be met by redoubled efforts to succeed. (*n* Counteractive Achievement)
4. I am clearly conscious of my main goals in life. (Ego Strength)
5. I usually can think of something to say in my own defense. (*n* Defendance)
6. I love good music and good wine and have become quite a connoisseur in such matters.
7. I usually decide for myself what I want to do and do it. (Conative Conjunctivity)
8. I am more apt to give in to an opponent than to keep on arguing with him. (Dejection Pessimism)
9. I spend a lot of time thinking about my shortcomings. (Ego-Ideal Intragression)
10. I enjoy organizing or directing the activities of a group—team, club, or committee. (*n* Dominance)
11. I like to amuse people by wisecracking or playing the fool. (*n* Recognition)
12. I feel that I am temperamentally different from most people. (Narcism)

self-ratings were confirmed by the clinical interviewer were asked to serve in the experiment.

The Psychological Insight Test. This test, consisting of two hundred statements, was constructed by the writer according to principles outlined by Murray and his co-workers (21). It provides a measure of the six personality variables within the Strong Ego syndrome and the six within the Weak Ego. Since three of the variables, *n* Recognition, Dejection-Pessimism, and Ego-Ideal Intragression, appear in both syndromes, only nine variables are included in the test. Each is measured by a set of 20 statements. Twenty additional statements which tap a variety of disparate needs were included but not scored. (See Table 1[12].)

The test, self-administered, requires *S* to rate himself on each statement on a +3 to −3 scale, the scale's zero point being the behavioral and emotional reactions of most men of *S*'s age, i.e., "the hypothetical average among college men." Thus, +3 represents *very much more often* than the average college man, −3 *very much less often*, +1 average but on the high side, −1 average but on the low side, etc.

When scoring the test, Strong Ego and Weak Ego scores are computed separately for each *S*. Each score consists of two parts: the algebraic sum of those variables which should be high for the given pattern, and the algebraic sum of those variables which should be low. A Strong Ego's score on Strong Ego variables, thus, would yield a +, − pattern, a Weak Ego's score on Weak Ego variables, a −, + pattern. Arithmetically, the scores on the Strong Ego variables can range from +240, −120 to −240, +120. The same is true for the Weak Ego variables.[13]

The Psychological Insight Test was administered to 140 male undergraduates enrolled in an introductory course in psychology at Harvard University under "take home" conditions. Eighty-six tests were returned. A Strong and a Weak Ego score was computed for each *S*, following the procedure described above.

In order to determine the range of Strong and of Weak Ego scores obtained by this sample, the bipart scores were combined into a single score by changing the sign of the second half of the score and adding the resulting sum to the first half of the score. For example, a score of +50, −25 would become +75, a score of −50, +25 would become −75. The scores of the eighty-six *S*s on the Strong Ego variables ranged from +174 to −174; on the Weak Ego variables from +152 to −112. The *S*s whose scores fell in the upper quartile of the Strong Ego distribution were selected as potential Strong Ego *S*s. The *S*s whose scores fell in the lower quartile of the Weak Ego distribution were selected as potential Weak Ego *S*s. Of these, twelve Strong Ego scoring *S*s and eleven Weak Ego scoring *S*s agreed to come for an interview to the Harvard Psychological Clinic for the purpose of "possible participation in some experiments at the Clinic."

The interview. The interview, conducted either by Dr. Henry A. Murray or by Dr. Robert W. White, was structured around the Strong and Weak Ego syndrome variables, definitions and criteria for which were provided to the interviewer. Using a 0–5 point scale, the interviewer rated *S* on eight of the nine variables in these syndromes[14] and on the basis of these ratings labeled *S* as a "Strong Ego", a "Weak Ego" or "neither consistently Strong nor consistently Weak."

If the self-ratings and the interviewer's ratings were in agreement, either consistently "Strong" or consistently "Weak," *S* was asked by the interviewer to serve in the experiment.[15] In this way, the two groups

[12] These statements were added both to disguise the nature of the test and to facilitate scoring. Nine variables and a "joker" variable repeated 20 times each allowed for the following scoring arrangement: statements measuring a given variable, for example *n* Recognition, were placed in the first, eleventh, twenty-first, etc., positions; a second variable, Narcism, in the second, twelfth, twenty-second, etc., positions. Since the statements for a given variable are sufficiently diversified in the test the scheme is not readily detected by *S*. This plan is demonstrated by the sample set of statements in Table 1. The need which each measured has been noted in parentheses following the statement.

[13] To obtain the maximum Strong Ego score, *S* would have to rate himself +3 on all statements (80) pertaining to *n* Recognition, *n* Dominance, Ego Strength, and Conative Conjunctivity, and −3 on all statements (40) pertaining to Dejection-Pessimism and Ego-Ideal Intragression. It is to be noted that unlike the McClelland *et al.* (16) *n* Achievement variable, the present syndromes do not represent opposite ends of the same personality continuum. Only some of the variables appear in both syndromes.

[14] Through an oversight on the part of the writer, one variable, Narcism, had been omitted from the interview schedule.

[15] A partial measure of the sensitivity of *The Psychological Insight Test* is reflected by the fact that only five of the 23 *S*s interviewed had to be rejected for the

of nine *Ss* each were selected. Their group membership "Strong" or "Weak" was not known to the writer until after the prediction scores had been computed at the end of the experiment.

Materials. The experimental materials were the same as those used in the earlier experiment (1). They consisted of two comparable sets of scrambled phrases to be arranged into meaningful sentences. Of the 12 sentences in each set, six were readily solvable within the two-minute time limit, three too difficult for control *Ss* to solve within the time limit, and three were unsolvable. Each solvable sentence permitted four alternative solutions.

Procedure. The procedure of the earlier experiment (1) was followed exactly. The atmosphere and instructions in Session I, Task Orientation, were informal and objectively non-self-esteem threatening since ostensibly *S* was merely helping *E* find out whether the materials would be suitable for a later experiment. The atmosphere and instructions in Session II, Ego Orientation, a week later, were formal and objectively threatening to self-esteem, sentence solution now being presented as "a brief intelligence test which the Clinic had devised for use by the Army in selecting officer training school candidates." To heighten the threat, moreover, a male accomplice, introduced as another *S*, worked in the same room on the same sentences and "successfully completed" all tasks within the time limits allowed.

At the end of the session, following a ten-minute interpolated projective test task,[16] incidental recall of the sentences was measured.

Treatment of recall data. In order to adjust for individual differences in number of sentences actually completed, two percentage recall scores were computed for each *S*: percentage of completed tasks recalled, and percentage of interrupted tasks recalled. This was done separately for each session. Since percentage scores are nonlinear, each was converted into angles by the arc $\sin\sqrt{\text{percentage}}$ transformation (27, Table 16.8, p. 449) before applying tests of significance.

Prediction of whether a given *S* fitted the Strong or the Weak category was based on his total recall score. The formula used for obtaining *S*'s total recall score was as follows:

Total Recall Score = $(In_I - C_I) - (In_{II} - C_{II})$

where In_I refers to *S*'s transformed score for the recall of interrupted tasks in Session I, C_I to his transformed score for the recall of completed tasks in Session I, In_{II} to his transformed score for the recall of interrupted tasks in Session II, and C_{II} to his transformed score for the recall of completed tasks in Session II.

To support the earlier results (2), the total recall score of a Strong Ego should be positive, a Weak Ego negative. The correctness of the prediction was tested by means of Fisher's (11) exact test.

To test whether the experimental groups differed in the direction of their recall within each orientation, the Mann-Whitney *U* Test (16) was applied to the transformed percentage recall scores.

Treatment of performance data. Differences between the two groups of *Ss* in number of sentences solved, in number of alternative solutions per sentence achieved, in time to achieve a first solution, and in total number of sentences recalled were analyzed by means of analysis of variance.

TABLE 2

PREDICTION OF STRONG AND WEAK EGO, SELECTIVE RECALL PATTERNS FOR TWO GROUPS OF MALE UNDERGRADUATES, A STRONG EGO GROUP AND A WEAK EGO GROUP

Category	Strong Ego Recall Pattern[a]	Weak Ego Recall Pattern[b]	P*
Expected	9	9	
Obtained	9	8	
Correctly Predicted	7	6	.045

[a] Based on total recall pattern over two sessions, $(In_I - C_I) - (In_{II} - C_{II})$; to be counted as a Strong Ego pattern the score must be positive.

[b] Based on total recall pattern over two sessions, $(In_I - C_I) - (In_{II} - C_{II})$; to be counted as a Weak Ego pattern the score must be negative.

* Computed by means of Fisher's (11) exact test.

Results

Prediction of group membership. Nine of the 18 *Ss* had positive total recall scores, eight negative total recall scores, one a zero score. Since a zero recall score does not permit prediction of group membership, this *S* was omitted from the prediction analysis.[17] As shown in Table 2, the prediction of Strong Ego was found to be correct for seven of the nine positive total recall scores, the prediction of Weak Ego for six of the eight negative ones. Fisher's exact test yields a P value of .045 for these data. The null hypothesis that the direction of selective recall is unrelated to personality structure, as measured in this experiment, can therefore be rejected.

Selective recall of Strong and Weak Ego Groups within each orientation. The analyses from this point on are based on the results obtained by the *Ss* within the two experimental groups, i.e., the nine Strong and the nine Weak Ego *Ss* as selected by the combined criteria of the self-rating scale and interview, described above.

Differences in percentage recall scores ob-

experiment. One *S* scored "Strong" on the self-rating scale and was rated "Weak" by the interviewer and four *Ss* did not consistently fit the interview criteria for either "Strong" or "Weak."

[16] The results of this test, *The Mind Reading Test*, are not included in this report.

[17] This score was subsequently identified as the score of a Strong Ego *S*.

TABLE 3
MEAN RECALL SCORES OF STRONG EGO AND WEAK EGO MALE SUBJECTS TESTED UNDER BOTH TASK-ORIENTED AND EGO-ORIENTED CONDITIONS*

	Task-Orientation		Ego-Orientation	
	In	C	In	C
Strong Ego *Ss*	.37	.57	.26	.46
Weak Ego *Ss*	.29	.48	.31	.32
	P†		*P*†	
In	.07		.25	
C	.27		.06	
In—C	.45		.01	

* Percentage recall scores converted into angles by the arc sin $\sqrt{}$ percentage transformation (27, Table 16.8, p. 449).
† Computed by means of Mann-Whitney *U* Test (14).

tained by the *Ss* in the two experimental groups are shown in Table 3. Under Task Orientation, the Strong Ego group tends to recall more interrupted tasks than the Weak Ego *Ss* ($P = .07$). The Weak Ego *Ss* recall more completed tasks than do the Strong Ego *Ss*, but the difference is not statistically significant ($P = .27$). There is also no difference in the so-called Zeigarnik effect, In − C ($P = .45$), though the difference is in the expected direction.

Under Ego Orientation, Weak Ego *Ss* recall more interrupted tasks than do Strong Egos, but the difference is not statistically significant. Strong Ego *Ss* tend to recall more completed tasks than do Weak Egos, and this difference approaches significance ($P = .06$). The difference in recall as measured by In − C for the two groups is significant ($P = .01$). In other words, the discrepancy in favor of the greater recall of completed than of interrupted tasks is greater for the Strong Ego *Ss* than for the Weak Ego *Ss*.

Performance of the Strong and Weak Ego Groups within each orientation. Tables 4, 5, and 6 summarize the group differences for the performance variables. Table 4 shows that under Task Orientation, the two groups are equally efficient at solving the sentences. Under Ego Orientation, both groups lose in efficiency, the Strong Ego *Ss* losing less than the Weak Ego *Ss*, but the differences are not significant.

Since each solvable sentence permitted four solutions, a more sensitive measure of performance loss from Task Orientation to Ego Orientation is afforded by the mean number of alternative solutions achieved under each condition. As Table 5 shows, the mean for Weak Ego *Ss* under Task Orientation is higher than for Strong Egos, but the difference is not significant. Under Ego Orientation, both groups show a drop. When differences *between* the two sessions are analyzed, the loss for the Weak Ego *Ss* is found to be significantly greater than for the Strong Ego *Ss* ($P = .05$).

The poorer performance of Weak Ego *Ss*, as the instructions become more ego threatening, may result from increased rigidity, i.e., having once found a solution they are less able than Strong Egos to disturb that structure and to attain an alternative solution. The possibility of greater rigidity in Weak Egos is supported by the data concerning time to achieve a first solution. These data, summarized in Table 6, show no significant difference between groups in time required to achieve a *first* solution. In other words, Weak Ego *Ss* are not slower to find an initial solution to the problem when the instructions become more ego threatening (Table 6), but are significantly less able to restructure and achieve additional solutions within the two-minute time limit per sentence (Table 5).

As Table 7 shows, there is no significant overall memory loss as the conditions become more threatening to self-esteem. The differences within each experimental session, moreover, do not even approach significance.

The selective recall differences reported above, therefore, represent qualitative differences as a function of ego strength, not quantitative differences in total memory. The performance data indicate both qualitative and quantitative group differences. Under objectively non-self-esteem threatening conditions, the performance of the two groups is the same. When self-esteem is objectively threatened, however, Weak Ego *Ss* show the greater loss in efficiency.

PART II

THE RELATIONSHIP BETWEEN EGO STRENGTH AND *N* ACHIEVEMENT

In order to fit the data of Part I into the results of recall studies done within the McClelland *et al.* (18) framework of achievement motivation, the relationships between ***n*** Achievement as a single parameter of per-

TABLE 4

ANALYSIS OF VARIANCE FOR NUMBER OF SENTENCES SOLVED BY STRONG AND WEAK EGO COLLEGE MALES UNDER TWO EXPERIMENTAL CONDITIONS

Task Orientation					Ego Orientation					Difference Between Orientations (Ego − Task)				
	df	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*
Total	17	10.3			Total	17	24.3			Total	17	27.1		
Between *Ss*	1	.1	.1	.16	Between *Ss*	1	4.5	4.5	3.66	Between *Ss*	1	5.5	5.5	4.07
Within	16	10.2	.64		Within	16	19.8	1.23		Within	16	21.6	1.35	
	Strong Egos		Weak Egos			Strong Egos		Weak Egos			Strong Egos		Weak Egos	
Mean	5.3		5.4		Mean	5.1		4.1		Mean	−0.2		−1.3	

TABLE 5

ANALYSIS OF VARIANCE FOR NUMBER OF ALTERNATIVE SOLUTIONS BY STRONG AND WEAK EGO COLLEGE MALES UNDER TWO EXPERIMENTAL CONDITIONS

Task Orientation					Ego Orientation					Difference Between Orientations (Ego − Task)				
	df	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*
Total	17	133.8			Total	17	73.8			Total	17	187.5		
Between *Ss*	1	8.1	8.1	1.03	Between *Ss*	1	5.6	5.6	1.3	Between *Ss*	1	44.5	44.5	5.0*
Within	16	125.7	7.9		Within	16	68.2	4.3		Within	16	143.0	8.9	
	Strong Egos		Weak Egos			Strong Egos		Weak Egos			Strong Egos		Weak Egos	
Mean	8.5		9.8		Mean	6.7		5.6		Mean	−1.8		−4.2	

* Indicates significance beyond the .05 level of confidence.

TABLE 6

ANALYSIS OF VARIANCE FOR AVERAGE TIME IN SECONDS TAKEN BY STRONG AND WEAK COLLEGE MALES TO ATTAIN A FIRST SOLUTION UNDER TWO EXPERIMENTAL CONDITIONS

Task Orientation					Ego Orientation					Difference Between Orientations (Ego − Task)				
	df	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*		*df*	Sum of Squares	Variance	*F*
Total	17	3766.54			Total	17	7212.72			Total	17	5304.37		
Between *Ss*	1	21.35	21.35	.01	Between *Ss*	1	90.60	90.60	.20	Between *Ss*	1	23.35	23.35	.07
Within	16	3745.19	234.07		Within	16	7122.12	455.13		Within	16	5281.02	330.06	
	Strong Egos		Weak Egos			Strong Egos		Weak Egos			Strong Egos		Weak Egos	
Mean	68.7		70.9		Mean	78.3		82.7		Mean	+9.6		+11.9	

sonality and the personality syndromes Ego Strength and Ego Weakness have to be examined. Since *n* Achievement was not measured in the prediction study, a new experiment was designed. The specific question it asks is: Do *Ss* who score high on *n* Achievement, as measured by the TAT method of McClelland *et al.* (18), more often fit only one of Alper's

TABLE 7

ANALYSIS OF VARIANCE FOR NUMBER OF SENTENCES RECALLED BY STRONG AND WEAK EGO COLLEGE MALES UNDER TWO EXPERIMENTAL CONDITIONS

Task Orientation					Ego Orientation					Difference Between Orientations (Ego − Task)				
	df	Sum of Squares	Var-iance	*F*		*df*	Sum of Squares	Var-iance	*F*		*df*	Sum of Squaree	Var-iance	*F*
Total	17	30.9			Total	17	54.9			Total	17	61.8		
Between *Ss*	1	2.7	2.7	1.5	Between *Ss*	1	1.3	1.3	.38	Between *Ss*	1	.2	.2	.05
Within	16	28.2	1.8		Within	16	53.6	3.4		Within	16	61.6	3.85	
	Strong Egos		Weak Egos			Strong Egos		Weak Egos			Strong Egos		Weak Egos	
Mean	5.4		4.7		Mean	4.2		3.7		Mean	−1.2		−1.0	

two personality syndromes, and *Ss* who score low on *n* Achievement, the other?

Although the *Ss* in this experiment were college females, rather than males, the sex factor should not affect the relationships being studied here since: (*a*) the design of the present experiment takes into account the conditions that should assure a high scoring *n* Achievement group; and (*b*) there are no reports of sex differences in the recall of interrupted vs. completed tasks. Indeed, in the Atkinson and Raphelson (6) study, the authors included the recall results of 22 female *Ss* "to add stability to the recall scores of men" (6, p. 352). They do not report separate recall results for the female sample.

Sex differences have been reported, however, in studies of *n* Achievement. McClelland *et al.* summarize them as follows: "(1) women get higher *n* Achievement scores than men under neutral conditions-two studies; (2) Achievement-Orientation does not increase *n* Achievement in women-three studies; (3) women's *n* Achievement scores seem as valid as men's in that they relate to performance in the same way; (4) the failure to find an increase in *n* Achievement score from achievement arousal is probably not due to the fact that the scores are already at a maximum" (18, p. 178). The clue for arousing higher *n* Achievement scores in women is seen by McClelland *et al.* to come from a study by Field (10) who reported that *n* Achievement in women is tied up with social acceptability rather than with reference to leadership capacity and intelligence. In Field's study, for example, females who were rated by their classmates as "liked" scored higher on *n* Achievement than those rated as "disliked," the ratings having been made after *n* Achievement had been measured and without advance knowledge that such ratings would be asked for. McClelland *et al.* conclude from these data that "if you want to arouse *n* Achievement in women, refer, as Field did, to their social acceptability; if you want to arouse *n* Achievement in men, refer, as we did to their leadership capacity and intelligence" (18, p. 181). Since Morrison (20) has subsequently confirmed Field's thesis by finding that female college leaders score higher on *n* Achievement than do female nonleaders, we may paraphrase McClelland *et al.*, as follows: if you want to be assured of a high *n* Achievement scoring female sample, use college leaders as *Ss*. Such students formed the experimental group in the present study.

An added advantage of a leader group for present purposes is the relationship reported by Atkinson and Raphelson (6) between selective recall and *n* Affiliation ("motivation for social acceptance" (6, p. 350)). Under a very relaxed orientation, *Ss* who scored high on *n* Affiliation, as measured by the TAT technique, recalled more interrupted than completed tasks, while *Ss* who scored low on *n* Affiliation showed the opposite recall pattern. Both Atkinson (5) and Atkinson and Raphelson (6) report that *Ss* who score high on *n* Achievement under Achievment Orientation, recall more interrupted tasks than do *Ss* who score low on *n* Achievement. These findings taken together suggest the possibility that *Ss* who score *high* on *n* Affiliation and *low* on *n* Achievement might show the recall pattern of Alper's

Strong Ego *Ss*; namely, greater recall of interrupted tasks under Task Orientation, of completed tasks under conditions which objectively stress that *S*'s performance is being measured. Similarly, *Ss* who score low on *n* Affiliation and high on *n* Achievement might be expected to show the reverse recall pattern characteristic of Alper's Weak Ego *Ss*. The fact that the recall of completed tasks has been found in the past to be a function of increased experimental stress, not of the strength of *S*'s *n* Achievement, may function against the hypothesized relationships between Ego Strength, *n* Affiliation, *n* Achievement, and recall pattern. On the other hand, since Atkinson and Raphelson (6) do not report how high *n* Affiliation *Ss* scored on *n* Achievement, though both measures were obtained for each *S*, nor have they tested the selective recall of the *same Ss* under both Task and Achievement Orientation, there is at present no clear contrary evidence against the hypothesis.

Method

The Ego Strength measure. The Psychological Insight Test, as described in Part I, was used to measure Ego Strength. Designed originally for use with college males, it was revised for female *Ss* only by changing "male" to "female," "men" to "women," "his" to "her," etc., in the directions and in the body of the scale. As in Part I, it was taken under "take home" conditions. In the instructions, *E* attempted to keep the conditions as nonthreatening to self-esteem as possible by emphasizing both that names would be removed after coding and that the results would be used as part of a study of college students in general, not as tests of the participating individuals.[18]

The algebraic scoring scheme, as described in Part I, was used to determine *S*'s score on each of the nine variables measured. To fit the Strong Ego pattern, *S*'s self-ratings would involve positive scores on *n* Recognition, *n* Dominance, Ego-Strength and Conative Conjunctivity, negative scores on Dejection-Pessimism and Ego-Ideal Intragression. To fit the Weak Ego pattern, *S*'s self-ratings would involve negative scores on *n* Recognition, *n* Defendance, *n* Counteractive Achievement and Narcism, positive scores on Dejection-Pessimism, and Ego-Ideal Intragression. The Strong Ego, thus, would have a +, − pattern of scores, the Weak Ego a −, + pattern. A +, + pattern can be regarded as an inconsistent Strong Ego pattern, a −, − pattern an inconsistent Weak Ego pattern.[19]

The n Achievement measure. The choice of picture stimuli for measuring *n* Achievement was influenced by: (*a*) the finding by Veroff *et al.* (28) that pictures of male characters yield higher *n* Achievement scores than pictures of female characters for both male and female college *Ss*; (*b*) the writer's clinical experience with the usefulness of particular TAT pictures for evoking achievement needs; and (*c*) an interest in measuring parent-child relationships as perceived by high and low scoring *Ss* on the *n* Achievement variable. These criteria resulted in the selection of four pictures: Cards 1 and 7GF from *The Thematic Apperception Test* (22), i.e., boy and violin, and young girl and mother-figure, respectively, and Cards 10G and 11G from *The Michigan Picture Test* (19), i.e., young girl and father-figure, and young girl in empty classroom. The *n* Achievement scores reported here are based only on two of these, the boy and violin, which are listed among the McClelland *et al.* (18) stimuli for measuring *n* Achievement, and the girl in classroom which can be considered a female counterpart of their Card H, "a boy in checked shirt at desk, an open book in front of him" (18, p. 375). Since the young girl is typically seen as about ten years old, as is also the case for the boy in the the boy and violin picture, when the storyteller mentions age at all, these two pictures are particularly well matched for tapping reactions to sex differences of the stimuli.

To facilitate group presentation of the pictures, permission was obtained from the respective publishers to make the necessary slides as well as certain changes in the stimuli. For example, in Card 11G, the background of the empty classroom was blurred, giving the background less "an empty classroom" appearance.

Experimental design. Three groups of college females were used as *Ss* in this experiment, an experimental group and two control groups. The experimental group was comprised of 28 upperclassmen, all of whom had been elected to high office by their peers. Achievement motivation was tested by a college senior, herself a college leader.[20] The conditions of testing and the instructions were casual, the intention being merely to task-orient these *Ss*. To this end, the group was told by the student *E* that "some studies of college students were being made and that they had been asked to help." A second group, 24 upperclassmen in an advanced course in psychology, served as a sample of upperclass nonleaders. Office holders in the group were eliminated from the sample. The test was administered by the instructor in the course immediately following the completion of an hour examination in the course in which these students were enrolled. The instructions emphasized that the test measured personality and that it would be scored by the instructor. This group, therefore, was an ego-oriented group and was tested immediately following a real-life achievement situation under ego-oriented conditions. The third group was

[18] It is quite likely that such instructions would be more ego-threatening to some, less to others. To label them nonthreatening to self-esteem, then, is to stress *E*'s ***intent***, not *S*'s involvement.

[19] Quantitative scores were obtained, as well as this pattern score, by following the scheme described in Part I. Only pattern scores are considered here since the concern is merely with presence or absence of a given self-rating pattern in relation to the *n* Achievement variable.

[20] The writer wishes to thank Miss Maud Hazeltine, President of Student Government at Wellesley College in 1955–56, for administering the test to the leader group.

comprised of 24 students enrolled in an introductory course in psychology. These *Ss*, largely sophomores, constituted another nonleader sample. Their achievement motivation was measured during a regular class hour under casual task-oriented conditions. The instructions emphasized that this was merely "a demonstration of a psychological technique." Since demonstrations were frequent in the course, this device was used for task-orienting these *Ss*.

In all three groups, the test was administered under group testing conditions. The McClelland *et al.* (18, p. 98) instructions and questions followed the special instructions, task- or ego-oriented, for the given group, as noted above.

The following predictions were made:

1. Following Morrison (20), the mean *n* Achievement score of the task-oriented leader group will be significantly higher than the mean *n* Achievement score of the ego-oriented nonleader group and of the task-oriented nonleader group;

2. Following Veroff *et al.* (28), the mean *n* Achievement score of the ego-oriented nonleader group will not differ significantly from the mean *n* Achievement score of the task-oriented nonleader group.

3. Following Atkinson (5) and Atkinson and Raphelson (6), in the leader group only, *Ss* whose self-ratings fit the Strong Ego pattern will more often score *low* on *n* Achievement than high on *n* Achievement, while *Ss* whose self-ratings fit the Weak Ego pattern will more often score *high* on *n* Achievement. The prediction was limited to the leader group since only this group meets the criterion of high *n* Affiliation required by the Atkinson and Raphelson (6) recall findings as well as Morrison's (20) *n* Achievement findings for female *Ss*.

Results

Scoring of Ego Strength and of *n* Achievement was done by the writer. Since the Ego Strength measure entails merely arithmetical addition of *Ss*' self-ratings, it presents no problem with respect to scorer reliability, a problem which arises, however, with the *n* Achievement measure. Two steps were taken to check on the reliability of the writer's scoring of *n* Achievement. First, several months after the initial self-instruction period with the McClelland *et al.* scoring categories (18, Chap. IV), the writer rescored the first story of each of the 30 *Ss* included in their Appendix I. A rank-order correlation test of the writer's scores checked against the scores reported for these stories by McClelland *et al.* (18) in their Appendix II yielded a rho value of +.85. Secondly, to check the writer's self-consistency for scoring *n* Achievement, she rescored the stories told by the leader group one month after her original scoring. The rho value here was +.96. Since these values of rho compare favorably with other reliability tests cited by McClelland *et al.* (18, pp. 185–187), it is assumed that the scoring is reliable.

TABLE 8

MEAN *n* ACHIEVEMENT SCORES OF COLLEGE FEMALES UNDER VARIOUS ORIENTATIONS (Stimuli Were TAT Card I and Michigan Picture Test Card 11G)

	College leaders, task-oriented	Nonleaders, ego-oriented	Nonleaders, task-oriented
Mean *n* Achievement Score	3.1	4.1	3.5

	t	*df*	*P*
College leaders — nonleaders, ego-oriented	−.85	50	.50 − .40
College leaders — nonleaders, task-oriented	−.34	50	.80 − .70
Nonleaders, ego-oriented — nonleaders, task-oriented	−.55	46	.60 − .50

Analysis of the total mean n Achievement scores based on stories told to two pictures (TAT Card 1 and Michigan Picture Test Card 11G). The mean *n* Achievement scores for the three groups of *Ss* are shown in Table 8. The data indicate that the three groups do not differ on *n* Achievement. The first prediction, that the leader group would score higher than the nonleader groups, is therefore not supported. In fact, on an absolute basis, the leader group scored lowest, the ego-oriented nonleader group scored highest, the task-oriented nonleader group in the middle, though none of the differences are significant.

Since the mean *n* Achievement scores of the two nonleader groups do not differ, the second prediction is supported. The fact that one group was ego-oriented during the testing and the other task-oriented, in other words, does not affect their *n* Achievement scores.

Analysis of the effect of the sex of the character in the picture stimulus on n Achievement scores. To check the possibility that the three groups may have been responding differently to the two pictures, separate *n* Achievement scores were computed for each picture. Dependent *t* tests, comparing the scores on the two pictures within a group, yielded the results summarized in Table 9. These data indicate that the two pictures are not equally effective in arousing *n* Achievement for all groups. The male picture arouses significantly more *n* Achieve-

ment in the leader group than does the female picture, the mean of the former being 2.6, of the latter 0.7 ($P < .05$). The differences are in the same direction for the ego-oriented nonleader group, but they are not significant ($P < .20$). For the task-oriented nonleader group there is a reversal: the female picture tends to yield a higher mean *n* Achievement score than the male picture ($P < .10$).

In view of these findings, the first two predictions were tested for each picture separately. The results, analyzed by means of independent *t* tests, are summarized in Table 10.

As is shown in the left half of Table 10, Prediction 1 is partly supported for the male picture: *n* Achievement *is* significantly higher for the leader group than for the task-oriented nonleader group ($t = 7.96$; $P < .001$), but there are no differences between the leader group and the ego-oriented nonleader group ($t = -.53$; $P < .60$).

Prediction 2 is *not* supported since the mean *n* Achievement score of the ego-oriented nonleader group *is* significantly higher than the mean *n* Achievement score of the task-oriented nonleader group ($t = 2.04$; $P = .05$).

The right half of Table 10 summarizes the results for the female picture. Again the leader group does not differ from the ego-oriented nonleader group ($t = -.56$; $P < .60$) but does differ from the task-oriented nonleader group. The female picture yields a significantly *lower* mean *n* Achievement score for the task-oriented leader group than for the task-oriented nonleader group ($t = -2.16$; $P = .05$). The two nonleader groups also differ on the female picture, the ego-oriented nonleader group scoring significantly *lower* on *n* Achievement than the task-oriented nonleader group ($t = -5.43$; $P < .001$).

In brief, these results indicate that: (*a*) the sex of the character in the picture stimulus does affect the strength of the *n* Achievement score of college females; (*b*) *n* Achievement is more likely to be aroused in a female *S* by a male than by a female character if *S* either values social acceptance or is being tested in an achievement atmosphere; (*c*) *n* Achievement is more likely to be aroused in a female *S* by a female than by a male character if *S* does not value social acceptance and she is being tested in a task-oriented atmosphere.

Relation between Ego-Strength patterns and n Achievement. The *Psychological Insight Test* scores were labeled Strong Ego, Weak Ego, Not Consistently Strong, or Not Consistently Weak according to whether the score fitted a +, −; a −, +; a +, +; or a −, − pattern, respectively, as described earlier. The frequency of these patterns within the leader and nonleader groups is shown in Table 11. The two nonleader groups were combined here since a chi-square analysis revealed them not to be different ($\chi^2 = .75$; $P < .90$).

In all groups, the Weak Ego pattern occurs rarely. The Strong Ego pattern occurs most frequently, but not more frequently than would be expected by chance alone when

TABLE 9

MEAN *n* ACHIEVEMENT SCORES OF COLLEGE FEMALES BASED ON STORIES TOLD TO A MALE AND TO A FEMALE PICTURE STIMULUS

Group	Male Picture	Female Picture	*t*	*df*	*P*
College leaders, task-oriented	2.6	.7	2.42	27	.05–.02
Nonleaders, ego-oriented	3.0	1.1	1.47	23	.20–.10
Nonleaders, task-oriented	1.3	2.3	−1.92	23	.10–.05

TABLE 10

MEAN ACHIEVEMENT SCORES OBTAINED BY COLLEGE FEMALES ON STORIES TOLD TO A MALE AND TO A FEMALE PICTURE STIMULUS

Group	Male Picture			Female Picture		
	t	*df*	*P*	*t*	*df*	*P*
College leaders, task-oriented—nonleaders, ego-oriented	−.53	50	.60 – .50	−.56	50	.60 – .50
College leaders, task-oriented—nonleaders, task-oriented	2.04	50	.05	−2.16	50	.05
Nonleaders, ego-oriented—nonleaders, task-oriented	7.96	46	< .001	−5.43	46	< .001

TABLE 11
PATTERNS OF PSYCHOLOGICAL INSIGHT TEST SCORES OBTAINED BY FEMALE COLLEGE LEADERS AND FEMALE COLLEGE NONLEADERS

Group	*N*	Psychological Insight Test Patterns			
		Strong Ego	Weak Ego	Not Consistently Strong	Not Consistently Weak
Leaders	28	16	2	4	6
Nonleaders	46	21	6	15	4

TABLE 12
RELATIONSHIP BETWEEN EGO STRENGTH AND *n* ACHIEVEMENT OF FEMALE COLLEGE STUDENTS
(Analysis by chi square[a])

	N	No. Scoring Above the *n* Achievement Mean of the Total Group	No. Scoring Below the *n* Achievement Mean of the Total Group	χ^2	P[b]
Strong Ego *Ss*	16	3	13	6.31	< .02
Weak Ego *Ss*	2	0	2	—	—

[a] The correction for continuity was applied to these data.
[b] $df = 1$.

Strong Ego frequency is compared with the frequency with which all other patterns combined occur.

Table 12 shows the number of Strong Ego *Ss* in the leader group whose *n* Achievement scores fell above the group's mean, and the number whose scores fell below the group's mean. Similar data are included for the Weak Ego *Ss*.

The Strong Ego data indicate that significantly more *Ss* in the leader group whose self-ratings fit the Strong Ego pattern score *below* the mean of the group on *n* Achievement ($P = .02$). Thus, the prediction that Strong Ego *Ss* in the leader group would more often score low on *n* Achievement, not high, is supported.

The prediction that Weak Ego *Ss* in the leader group would score high on *n* Achievement was not tested since the pattern appeared so infrequently in this group.

A prediction of how Strong Ego *Ss* in nonleader groups would score with respect to their group means was not made since the *n* Achievement literature provided no basis for making a prediction. The fact that in both nonleader groups as many Strong Ego *Ss* scored above their group mean as below it, however, strengthens the importance of the the inverse relationship between *n* Achievement and Ego Strength found in the leader group data.

DISCUSSION

The results reported in Part I of this paper support the position taken earlier by the writer (1, 2, 3) that selective recall is lawfully related to the personality structure of the recaller. In the present study, two personality patterns, the Strong Ego and the Weak Ego, as previously (2) described, were studied further by selecting in advance of the experiment *Ss* who fitted these patterns. All *Ss* were tested under two experimental conditions, Task Orientation first, followed by Ego Orientation a week later. It was found that correct prediction of group membership on the basis of the selective recall pattern alone was possible significantly beyond chance: the Strong Ego pattern consisting of greater recall of completed than of interrupted tasks as the experimental conditions become more threatening to self-esteem; the Weak Ego pattern, of greater recall of interrupted than of completed tasks.

As in previous papers (2, 3), these results are interpreted to mean that Strong and Weak Ego *Ss* protect self-esteem in different ways. In situations in which completion is likely to be experienced as personal success, interruption as personal failure, i.e., when tasks are being performed under ego-oriented conditions, the Strong Ego is not overwhelmed by the failure, the Weak Ego is. The Strong Ego supports his self-esteem by focusing on the tasks he *did* complete, i.e., recalls more completed than interrupted tasks. This mechanism of defense is not available to the Weak Ego. Instead, the Weak Ego utilizes an ego-offensive mechanism (2): he focuses on his failures, i.e., recalls more interrupted than completed tasks, admitting his failures before others face him with having failed. Both types of recall, then, serve to reduce ego tensions under conditions experimentally designed to arouse ego, rather than task, tensions.

Similar recall patterns, but interpreted in very different ways, have been reported recently by Atkinson (5) and Atkinson and

Raphelson (6). The data of Part II support a reinterpretation of their findings. As a result, the *n* Achievement and Ego Strength studies can be shown to have yielded uniform patterns of selective recall.

Atkinson and Raphelson (6) have reported that *S*s who score high on *n* Affiliation (motivation for social acceptance) as measured by the TAT technique of McClelland *et al.* (18) under the "very relaxed" orientation, recall more interrupted tasks than do *S*s who score low on *n* Affiliation. As the conditions become more achievement oriented, the direction of selective recall correlates with *n* Achievement, not with *n* Affiliation. The *S*s who score high on *n* Achievement tend to recall more interrupted than completed tasks, while *S*s who score low on *n* Achievement recall more completed than interrupted tasks (5, 6). They interpret these results to mean that the *S* who recalls more interrupted tasks under relaxed conditions is reacting merely with task tensions. He wants to *please* the *E* and hence, like Zeigarnik's (29) *S*s, when a task tension cannot be resolved, because *E* has interrupted *S* on the task, the tension persists, making interrupted tasks more available for recall than completed ones. The greater recall of interrupted than of completed tasks by high *n* Achievers, and the reverse pattern by low *n* Achievers, under achievement-oriented conditions are explained as follows: high *n* Achievers have a "positive disposition to excel—are predominantly success-oriented and have as their goal the feeling of satisfaction accompanying personal accomplishment while persons lower in *n* Achievement are more concerned with avoiding feelings of failure" (5, p. 386), i.e., "lower *n* Achievement scores imply relatively greater anxiety about failure" (5, p. 389).

Thus, recalling more of one's failures (interrupted tasks) under achievement-oriented conditions, the pattern of the *S* who scores high on *n* Achievement, is interpreted by Atkinson to express *S*'s "positive disposition to excel," by Alper to mean that the Weak Ego, overwhelmed by the experience of failure, protects self-esteem by taking the offensive; he admits his failures before others accuse him of having failed. Recalling more of one's successes (completed tasks) under achievement-oriented conditions, the pattern of the low *n* Achiever, is interpreted by Atkinson (5) as an expression of anxiety over failure and by Alper as a way of supporting self-esteem through focusing on tasks one did complete. Seemingly opposite defense patterns, in other words, are seen to be operating to produce the same recall patterns.

The results in Part II of this paper provide a basis for showing how the two interpretations may be brought into a common theoretical framework. By obtaining both Ego Strength and *n* Achievement measures on female college leaders, a group which according to previous studies should score high on *n* Achievement (10, 20), it was found that *S*s whose self-ratings fit the Strong Ego pattern more often score *below* the group mean on *n* Achievement than above it. No relationship was found between *n* Achievement and the Strong Ego self-rating pattern in two control groups. Since, by definition, the leader group should value social acceptance, i.e., score high on *n* Affiliation, we now can relate the Atkinson (5) and Atkinson and Raphelson (6) findings to the Strong Ego selective recall patterns isolated by the writer. The greater recall of interrupted tasks under the Relaxed Orientation by *S*s who score high on *n* Affiliation (6), and of completed rather than interrupted tasks under conditions which objectively stress achievement by *S*s who score *low* on *n* Achievement (5, 6), is exactly the recall pattern expected of a Strong Ego *S*.

While the relationship between high *n* Achievement and the Weak Ego self-rating pattern could not be adequately tested in Part II because so few *S*s in the leader group fitted the Weak Ego pattern, the inverse relationship between the Strong Ego self-rating pattern and *n* Achievement permits at least a partial reinterpretation of the role *n* Achievement plays in selective recall. The results of Part II do not support Atkinson's (5, p. 389) proposition that "lower *n* Achievement scores imply relatively greater anxiety about failure" if low *n* Achievement is combined with the personality structure of a Strong Ego *S*. On the contrary, these *S*s are able to counteract the objective failure threat (the Strong Ego *S*s in Part I obtained significantly more alternative solutions than did Weak Ego *S*s). Counteraction suggests positive, constructive forces at work rather

than a mechanism of "anxiety-avoidance." In other words, it may be the high *n* Achievers as well as Zeigarnik's (29) "ambitious" *S*s, not the low *n* Achievers, who have the relatively greater anxiety about failure.

Further support for these interpretations comes from an analysis of the self-ratings of the leader group on the nine individual personality variables included in *The Psychological Insight Test*. As a group, these socially accepted *S*s significantly more often rated themselves high rather than low, i.e., above the average for college women, not below the average, on *n* Counteractive Achievement, Conative Conjunctivity and Ego Strength, and low rather than high on *n* Recognition, Dejection-Pessimism and Ego-Ideal Intragression. The variables on which they score equally often at the high or the low end are Narcism, *n* Defendance, and *n* Dominance. In this group, the *S*s accept as characteristic of themselves, then, only those variables which could be regarded as having a positive value in our culture. This finding is consistent with Martire's (17) finding that low *n* Achievers are less dissatisfied with themselves, less in "self-conflict," than are high *n* Achievers.

The low *n* Achiever who is a Strong Ego, in other words, is low on achievement motivation as long as the *objective* situation is unrelated to achievement. His Achievement needs are *not constantly* under tension. Yet when achievement is realistically called for, and failure is a realistic possibility, the Strong Ego can and does counteract it. The counteraction shows itself both in his selective recall and in his performance data. It should not be surprising, therefore, to find significant differences in selective recall between high and low *n* Achievers consistently reported only for Achievement Orientation (5, 6), while performance studies of high and low *n* Achievers continue to yield inconsistent results (6, 7, 12, 27). In some of these studies, high *n* Achievers perform better than low *n* Achievers under both Task Orientation and Achievement Orientation, in others only under Achievement Orientation. To understand these results better, it may be necessary to take into account both *n* Achievement and *n* Counteractive Achievement, to differentiate, for example, between high *n* Achievers who are high on *n* Counteractive Achievement and high *n* Achievers who are low on *n* Counteractive Achievement. Moreover, while low *n* Achievement and high *n* Counteractive Achievement appears to be the Strong Ego pattern, theoretically, Weak Ego *S*s could fall either into a low *n* Achievement, low *n* Counteractive Achievement or high *n* Achievement, low *n* Counteractive Achievement group. The two Weak Ego *S*s in the present leader group, for example, would fit the first pattern since they rated themselves low on *n* Counteractive Achievement and scored low on *n* Achievement as measured by the TAT technique.

While the relationships between low *n* Achievement and Strong Ego self-ratings hold statistically, and thus warrant the reinterpretation of the selective recall findings of studies done within the *n* Achievement framework, certain problems in the data of the present study, and certain objections which can be raised concerning them, need to be considered. The first problem centers around a difficulty raised in Part I. The recall patterns of Strong and Weak Ego *S*s were not statistically different under Task Orientation, though the overall differences between the sessions were significant. The earlier study (2) had indicated that both orientations would yield significant recall differences within sessions. The absence of significant recall differences under Task Orientation is not peculiar to the present study (see 3, 6, 13, 26). While no general explanation is available, a factor that may have been important in the present study may also have affected other studies. In the experiment reported in Part I, the experimental atmosphere may not have been as nonthreatening to self-esteem as *E* had intended, even though the instructions were the same as those previously used. In the earlier study (2), the *S*s were familiar with the clinic setting and with the clinic personnel. They came to the clinic many times since all were participating in an extensive personality study. In the present experiment, the *S*s served only in the one experiment. If this explanation has validity, it suggests that at a given point in the ego-stress continuum, ego tensions begin to function for Strong Ego *S*s also, and at this point there are no differences in the recall patterns of the two groups of *S*s. At a further point along the continuum, the recall patterns would again differ.

A second problem in relating the recall patterns emerging in the *n* Achievement studies to Alper's patterns, stems from possible differences in an "Achievement" vs. an "ego oriented" or "self-esteem threatening" atmosphere. While the terms themselves do not necessarily indicate important motivational differences, the writer would agree that so long as the conditions used to arouse the motive center around achieving, i.e., making a good record for oneself against some objective standard of excellence, the term "Achievement Orientation" expresses very precisely the motivation the *E* intends to have aroused. Yet in both the Atkinson and in the Alper studies, the arousal of more than achievement needs seems to have been intended since in both, the *S* is prevented from doing well on (i.e., completing) "tests." In the Atkinson study, *S* was told that the tests were "developed by psychologists during the war yeras in order to select people with high executive capacity, intellectual alertness, the capacity for making quick decisions, and leadership" (5, p. 382). In Alper's study, the tests were presented as "a brief intelligence test which the Army was finding useful in selecting candidates for Officer Training School" (2, p. 106). In other words, in both studies it was intended that *S* experience failure and that self-esteem be threatened. It is quite possible, of course, that Alper's instructions would be more threatening to draft age males than Atkinson's, if only because the incompletion occurred under more public conditions in the one than in the other; in Alper's study, an accomplice, of college age, "succeeded" on tasks which *S* was experimentally doomed to fail. If these differences affect recall, however, theoretically they should sharpen rather than alter the direction of recall differences.

A third problem arises with respect to whether the correlation between self-rating scores and *n* Achievement scores (the procedure followed in Part II) should be taken at face value since recent studies of the relationship between self-rating scores on achievement (*v* Achievement) and TAT measures of achievement motivation (*n* Achievement) are not entirely consistent. Two studies (8, 9) have used male *S*s, one (20) has used female *S*s. For male *S*s, de Charms *et al.* (9) report a positive correlation between *v* Achievement and *n* Achievement, the *P* value they obtained being $<.05$. But Child *et al.* (8) found no relation between *v* Achievement and *n* Achievement in their study of college males. For a female sample of college leaders, Morrison (20) reports that *v* Achievement and *n* Achievement are not correlated, while in the present study of female college leaders, there was a significant *inverse* relation between self-rating scores and *n* Achievement for the Strong Ego *S*s ($P < .02$).

The only other data available for comparing the relationships between *v* Achievement and *n* Achievement center on the F-scale scores on males in the de Charms *et al.* study, and on females in the present study. For males, de Charms *et al.* (9, p. 416) report that *v* Achievement correlates positively with total score on the F scale, while *n* Achievement does not. In the present study, *n* Achievement was also not related to F-scale scores, but self-rating scores were: Strong Ego *S*s tended to score *below* the group mean on the F scale ($P < .10$), while *S*s who did not fit the Strong Ego pattern, treated as a group, significantly scored *above* the group mean on the F scale ($P < .05$). In other words, it is the non-Strong Ego *S*s who are the more conformist, the more rigid *S*s among these college leaders, including the two *S*s whose self-rating scores fitted the Weak Ego pattern. The interpretation in Part I, that Strong Ego *S*s were less rigid than Weak Ego *S*s, as measured by their ability to obtain more alternative solutions in the scrambled phrases task, is consistent with the F scale findings in Part II and with the data in the Child *et al.* (8) study.

Since many variables have been shown to influence the strength of the *n* Achievement variable, including the factor of the sex of the picture stimulus, further experimentation is necessary before the differences in findings of various studies can be clearly understood. The specific pictures used, the experimental atmosphere in which *n* Achievement is measured, even the questionnaire items[21] used in the self-rating measurements have varied greatly

[21] In all studies, the items have been derived from the Murray *et al.* (21) approach to measuring need strength by the questionnaire method. The length of the scales and the specific items covered have differed, however.

from one study to another. On the other hand, the relationships among these variables revealed in the present study are consistent with the writer's interpretation that Strong Ego *Ss* score *low* on *n* Achievement essentially because their achievement needs are *not* under tension rather than because of concern "with avoiding feelings of failure" (5, p. 386). This interpretation is an essential link for relating the selective recall pattern of the low *n* Achiever in the Atkinson (5) and Atkinson and Raphelson (6) studies to Alper's Strong Ego pattern and thus for reaffirming both the theoretical framework presented earlier by the writer (2) and its applicability to previously reported contradictory or equivocal findings of other studies (3).

SUMMARY

Two experiments are reported here. The first was designed to validate the two major recall patterns that had emerged in a previous study of the relationship between personality structure and the recall of interrupted vs. completed tasks, the Strong Ego pattern and the Weak Ego pattern (2). For this purpose, two groups of college males, a Strong Ego and a Weak Ego group, were selected in advance of the experiment by means of a specially constructed self-rating scale and a clinical interview. Otherwise, the design of the experiment, procedure, and materials were the same as in the previous study. Using only the total recall score pattern as a measure of the change in direction of recall from a task-oriented to an ego-oriented session, it was possible to predict correctly the group membership of these *Ss*, as Strong or as Weak Ego *Ss*, significantly beyond chance. These results confirm, therefore, the earlier finding that Strong Ego *Ss* recall significantly more completed tasks as the experimental conditions become more threatening to self-esteem; Weak Ego *Ss*, interrupted tasks. In the present study, the recall patterns of the two groups did not differ significantly within the task-oriented session, but did differ within the ego-oriented session.

The second experiment was designed to test the relationships between these recall patterns and the *n* Achievement findings reported by Atkinson (5) and Atkinson and Raphelson (6). These authors have reported that high *n* Achievers, as measured by the TAT method of McClelland *et al.* (18), recall more interrupted than completed tasks as the experimental conditions become more clearly achievement-oriented, while low *n* Achievers show the opposite recall pattern.

Three groups of female *Ss* were used, and both *n* Achievement and Ego Strength were measured. One group was made up only of college leaders who had been elected to high office by their peers. Following Morrison (20), these *Ss* should be a high *n* Achievement scoring group and could serve, therefore, as the experimental group. Two other nonleader groups served as controls. A significant *inverse* relationship between Strong Ego self-rating scores and *n* Achievement was obtained only for the leader group. The two measures were not correlated in the control groups. These findings are used to show that the recall pattern of the low *n* Achiever, as reported by Atkinson (5) and Atkinson and Raphelson (6), corresponds to Alper's Strong Ego pattern. The usual interpretation of a low *n* Achievement score as implying "anxiety over failure" is re-examined in the light of these findings.

It cannot be concluded from these data, however, that the selective recall pattern of a high *n* Achiever would necessarily fit Alper's Weak Ego pattern. Not only were there too few Weak Ego *Ss* in the college leader sample to test the relationship between Ego weakness and *n* Achievement, but also, and of more importance from the point of view of theory, Ego Strength and Ego Weakness, as measured in this study, are not at opposite ends of a single personality continuum. We cannot assume, therefore, that Ego Strength and Ego Weakness would function with respect to *n* Achievement in diametrically opposite ways. Yet given a high *n* Achiever whose self-ratings are congruent with the Weak Ego personality pattern, including low *n* Counteractive Achievement, it would be consistent with the results of these experiments to expect him to show the selective recall and performance patterns characteristic of Weak Ego *Ss*.

The theoretical framework presented earlier by the writer (2) is thus reaffirmed. The usefulness of this framework for reconciling seemingly contradictory selective recall and performance data reported in other studies, as previously discussed (3), is also reaffirmed.

Specific findings concerning *n* Achievement in females, including the importance of the sex of the picture stimulus used and the experimental atmosphere in which *n* Achievement is measured, are discussed.

REFERENCES

1. ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.*, 1946, **41**, 403–420.
2. ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: correlation between experimental and personality data. *J. Pers.*, 1948, **17**, 104–137.
3. ALPER, THELMA G. The interrupted task method in studies of selective recall: a reevaluation of some recent experiments. *Psychol. Rev.*, 1952, **59**, 71–88.
4. ALPER, THELMA G., & BLACK, ADELAIDE A. The effect of instruction, task and population sample on mental set. *Amer. J. Psychol.*, 1949, **62**, 295–299.
5. ATKINSON, J. W. The achievement motive and recall of interrupted and completed tasks. *J. exp. Psychol.*, 1953, **46**, 381–390.
6. ATKINSON, J. W., & RAPHELSON, A. C. Individual differences in motivation and behavior in particular situations. *J. Pers.*, 1956, **24**, 349–363.
7. BROWN, R. W. A determinant of the relationship between rigidity and authoritarianism. *J. abnorm. soc. Psychol.*, 1953, **48**, 469–476.
8. CHILD, IRVIN L., FRANK, KITTY F., & STORM, T. Self-ratings and TAT: their relations to each other and to childhood background. *J. Pers.*, 1956, **25**, 96–114.
9. DECHARMS, R., MORRISON, H. W., REITMAN, W., & MCCLELLAND, D. C. (Eds.) *Studies in motivation*. New York: Appleton-Century-Crofts, 1955. Pp. 414–427.
10. FIELD, W. F. The effects on thematic apperception of certain experimentally aroused needs. Unpublished doctor's dissertation, Univer. of Maryland, 1951.
11. FISHER, R. A. *Statistical methods for research workers*. (6th ed.) Edinburgh: Oliver and Boyd, 1936.
12. FRENCH, ELIZABETH G. Some characteristics of achievement motivations. *J. exp. Psychol.*, 1955, **50**, 232–236.
13. GLIXMAN, A. F. An analysis of the use of the interruption-technique in experimental studies of repression. *Psychol. Bull.*, 1948, **45**, 491–506.
14. HORN, D. A study of personality syndromes. *Charact. and Pers.*, 1944, **12**, 257–274.
15. LEWIS, HELEN B. An experimental study of the role of the ego in work. I. The role of the ego in co-operative work. *J. exp. Psychol.*, 1944, **34**, 113–126.
16. MANN, H. B., & WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 1947, **18**, 50–60.
17. MARTIRE, J. G. Relationships between self-concepts and differences in strength and generality of achievement motivation. *J. Pers.*, 1956, **24**, 364–375.
18. MCCLELLAND, D., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.
19. The Michigan Picture Test. Chicago, Ill.: Science Research Associates, 1953.
20. MORRISON, H. W. Validity and behavioral correlates of female need for achievement. Unpublished master's thesis, Wesleyan Univer., 1954. In D. C. McClelland (Ed.), *Studies in motivation*. New York: Appleton-Century-Crofts, 1955. P. 422.
21. MURRAY, H. A. *et al. Explorations in personality*. New York: Oxford Univer. Press, 1938.
22. MURRAY, H. A. *Thematic Apperception Test*. Cambridge, Mass.: Harvard Univer. Press, 1943.
23. ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reactions. *J. exp. Psychol.*, 1943, **32**, 64–94.
24. ROSENZWEIG, S., & MASON, GWENDOLYN. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, **24**, 247–265.
25. SANFORD, N., & RISSER, J. What are the conditions of defensive forgetting? *J. Pers.*, 1948, **17**, 244–260.
26. SEARS, R. R. Personality. In C. P. Stone, & D. W. Taylor (Eds.), *Annual review of psychology*. Stanford: Annual Reviews, 1950, Vol. I.
27. SNEDECOR, G. W. *Statistical methods*. (4th ed.) Ames, Iowa: Iowa State College Press, 1946.
28. VEROFF, J., WILCOX, SUE, & ATKINSON, J. W. The achievement motive in high school and college age women. *J. abnorm. soc. Psychol.*, 1953, **48**, 108–119.
29. ZEIGARNIK, BLUMA. On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of Gestalt psychology*. New York: Harcourt Brace, 1938. Pp. 300–314.

Received August 20, 1956.

THE RELATIONSHIP BETWEEN ADVANCED AGE AND MENTAL ABILITIES[1]

CHARLES R. STROTHER, K. WARNER SCHAIE, AND PAUL HORST

University of Washington

IN RESEARCH on aging, longitudinal studies in which the same population is followed over a long span of years are obviously to be preferred. There are an increasing number of such studies covering the periods of childhood, adolescence, early and middle adulthood (1, 2, 5), but a paucity of longitudinal data is as yet available on older age groups. The principal difficulties involved in such studies are the long period of time required for completion and the unavoidable attrition as the original subjects advance in age, due to such factors as mobility, illness, and death.

Cross-sectional studies, on the other hand, in which population samples in different age groups are compared, are open to several criticisms (3, 10). One criticism is concerned with the differences in educational background that are likely to characterize different age groups, usually to the advantage of the younger group. A second criticism is directed toward differences between age groups with respect to range of talent. For various reasons, it is difficult to obtain comparably broad samples of older and younger populations. Older subjects are more frequently drawn from institutions for the aged or from small groups of volunteers, while the younger subjects represent broader segments of the population. There are also differences in motivation and in attitudes toward, and familiarity with, psychological tests, which may tend to favor the younger groups. Finally, while the disabilities of older people with respect to speeded tests have been well established, few attempts have been made to apply appropriate corrections or to employ tests where speed is not a factor.

While these difficulties are serious, they may be minimized by restricting the groups compared to a fairly narrow age range. In the present study, the age range has been limited to the period 70–88 years. This restriction made possible the selection of subjects who were relatively homogeneous with respect to educational background, socioeconomic status, motivation, and familiarity with psychological tests. In this restricted population, differences in "primary mental abilities," "memory," "psycho-motor speed," and "motor rigidity" were determined by appropriate tests. Correction factors were then applied to the mental abilities scores to control the effects of loss of memory, reduced speed, and increased rigidity.

[1] This investigation was part of an interdisciplinary research project at the University of Washington, supported by a grant from the Agnes H. Anderson fund. It was reported at the 1956 meeting of the American Psychological Association. The authors wish to express their appreciation to Mary Marjorie Campbell who assisted in collection of the data.

METHOD

Subjects

All subjects (*Ss*) were volunteers obtained from the membership of a retired university faculty group or from among retired academic or professional workers who had responded to appeals in newspapers published in the city in which the university was located. The *Ss* were interviewed before being accepted for the research project, at which time it was explained that about eight hours, spread over several days, would be required for the physical examinations, psychological tests, and interviews involved in the larger project of which the present study is a part. At this point, the *S* was given an opportunity to withdraw from the study if he wished. It seems reasonable to conclude that the *Ss* who participated were at least as highly motivated as the average volunteer group.

The group consisted of 25 male and 25 female college graduates, ranging in age from 70 to 88 years, with a mean age of 76.5 years. Clinical and laboratory examinations indicated that all members of the group were in fair to superior physical condition for their age. No psychiatric complaints or symptoms were noted by the examining physicians.

Tests

In order to obtain estimates on more specific aspects of mental ability than would be yielded by a global score, Thurstone's Primary Mental Abilities Test (Intermediate Form) (12), which provides separate scores for Verbal-meaning, Space, Reasoning, Number, and Word-fluency factors, was selected. Previous work (6) had shown that the reliability and the range of difficulty for older subjects were adequate.

Since age changes in speed, memory, and rigidity are known to occur and might be expected to have some relationship to performance on the Primary Mental Abilities Test, measures of these variables were obtained. The Wechsler Memory Test (13) was selected as a measure of memory loss and Schaie's Test of Behavioral Rigidity (9) was used to obtain measures of "psycho-motor speed" and "motor rigidity."

All tests were administered as individual tests and

the order of administration was randomized in order to avoid systematic effects of fatigue.

RESULTS

The subjects were rank ordered by chronological age and divided into five groups of five males and five females each. A comparison by χ^2 showed no significant differences in years of education among the five groups. Standard Primary Mental Ability (PMA) scores with means of 50 and sigmas of 10 were then computed, by the use of Thurstone's conversion chart, from the norms for his 17-year-old group. Comparable standard scores for the speed and rigidity measures were computed, using Schaie's tables for his younger groups, ages 18–26. Standard scores on the Memory scale were computed from Wechsler's norms for his 20–30-year-old group (13).

Table 1 gives the mean standard scores for each of the five groups on all test variables. Examination of this table indicates that, for the 71-year-old group, the Reasoning and Spatial scores are below the mean of the 17-year-old reference population. If it can be assumed, on the basis of the amount of education and of previous occupational level, that the present group at an earlier age would have been above the mean of the general population on these tests, then an unspecified amount of loss in reasoning and spatial abilities has occurred prior to age 71. The Verbal-meaning, Number, and IQ scores are somewhat above the means for the younger reference population. Whether any loss has occurred on the Verbal and Number factors cannot be determined from these data. The relatively low scores of the 71-year-olds on the Word-fluency factor is most probably an artifact of sampling. The memory, speed, and motor-rigidity scores fall below the means for their respective reference groups.

TABLE 1
MEAN T SCORES FOR SUBGROUPS
(Young Adult Norms, M = 50, σ = 10)

Group	Mean Age	V*	S	R	N	W	IQ	Sp.	Mem	M-Rig
1	71	58.5	41.9	46.4	57.9	47.6	54.3	44.2	39.4	46.4
2	73	57.1	37.6	44.1	55.6	54.0	51.6	44.2	43.6	46.8
3	75	59.4	38.2	40.6	58.1	60.1	54.1	43.8	41.8	45.7
4	79	52.8	35.9	39.8	53.4	48.3	45.4	41.6	34.1	41.0
5	84	45.8	33.5	32.0	48.0	48.3	40.5	33.2	29.0	37.9
D_{1-5}		12.7	8.4	14.4	9.9	−0.7	13.8	11.0	10.4	8.5

* Key To Symbols:
V = Verbal Meaning
S = Space
R = Reasoning
N = Number
W = Word Fluency
IQ = Intelligence Quotient
Sp = Speed
Mem = Memory
M-Rig = Motor Rigidity

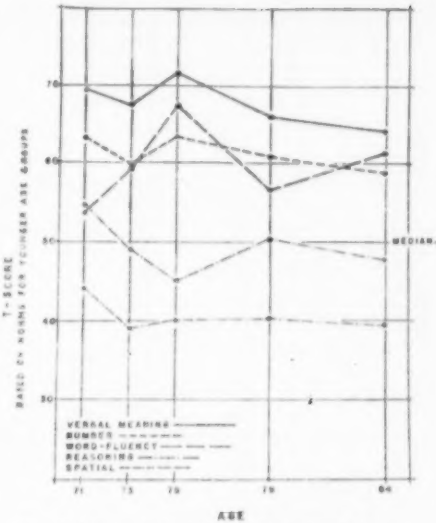

FIG. 1. MEAN *T* SCORES ON PRIMARY MENTAL ABILITIES FOR SUCCESSIVE AGE GROUPS

Figure 1 presents graphically the subgroup means for the PMA scores. For the Spatial and Reasoning factors, the scores show a steady decline after age 71. The Verbal-meaning, Number, and Word-fluency scores remain fairly high in the 71-year-old group, and then decline.

Figure 2 shows the curves of decline for the speed, memory, and motor-rigidity factors. There is little difference between the scores for the 71- and 73-year-old groups, but beyond age 73 the curves drop steadily. The most rapid loss occurs in memory.

Statistical analysis, using the *t* test to evaluate the significance of the difference between the youngest and the oldest groups, shows a decline significant at the .05 level on all factors except for Word-fluency.

The next step in the analysis was to determine whether the decline in PMA scores would be affected if corrections were made for the effects of lowered speed, increased rigidity, and memory loss. The previously described independent measures of these abilities were therefore correlated with the PMA scores. Intercorrelations among all these measures were computed on the IBM Type 604 electronic calculating punch, using the standard

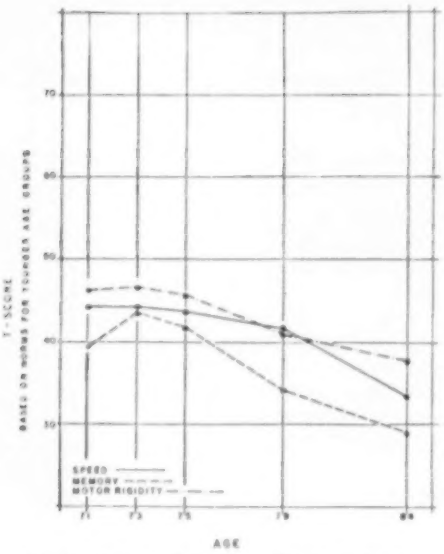

FIG. 2. MEAN *T* SCORES FOR MEMORY, SPEED AND MOTOR RIGIDITY

scores method (7). These intercorrelations, most of which are significant, are given in Table 2. Formulas were derived to correct for the effects of these factors on the PMA scores.[2] The correction formulas provide estimates of those parts of the PMA scores which are independent of the combined influence of the speed, rigidity, and memory scores, and yield corrected PMA measures in terms of *T* scores for the norm group.

It was not necessary to compute individual corrected PMA scores for the members of the experimental group in order to get estimates of the corrected subgroup means. These could be computed directly from the subgroup means given in Table 1 and the intercorrelations in Table 2, but the computations required the calculation of the beta regression vector for each PMA variable on the correction variables and also its multiple correlation with them.

[2] Materials presenting the mathematical derivation of the correction formulas, computational instructions, and tables of intermediate computations have been deposited with the American Documentation Institute. Order Document No. 5299 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.75 for microfilm or $2.50 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

TABLE 2
TEST INTERCORRELATIONS

	S	R	N	W	IQ	Sp.	Mem.	M-Rig.
Verbal Meaning	.348	.612	.300	.172	.630	.676	.375	.501
Space		.361	.193	.093	.470	.145	.266	.263
Reasoning			.395	.354	.757	.674	.571	.462
Number				.287	.784	.411	.318	.325
Word-fluency					.570	.396	.452	.025
Composite-IQ						.667	.554	.445
Speed							.410	.318
Memory								.406
Motor-Rigidity								

TABLE 3
BETA REGRESSION WEIGHTS AND MULTIPLE CORRELATIONS

	V	S	R	N	W	I.Q.
Speed	.571	.010	.495	.305	.298	.497
Memory	.013	.187	.292	.120	.429	.280
Motor Rigidity	.315	.184	.186	.180	−.244	.174
Σ	.899	.381	.973	.605	.483	.951
R^2	.548686	.099584	.586294	.222015	.305816	.564049
R	.741	.316	.766	.471	.553	.751

TABLE 4
MEAN T SCORES AFTER CORRECTIONS FOR THE EFFECT OF SPEED, MEMORY LOSS AND INCREASED MOTOR RIGIDITY

Mean Age	V	S	R	N	W	I.Q.
71	69.5	44.4	54.7	63.1	53.6	66.3
73	67.2	38.9	49.1	59.9	59.2	60.4
75	71.5	40.1	45.1	63.3	67.4	65.5
79	65.9	40.2	50.4	60.7	56.6	58.4
84	64.1	39.3	47.9	58.9	61.2	60.4
D_{1-5}	5.4	5.1	6.8	4.2	−7.6	5.9

The regression weights and multiple correlations are given in Table 3, and the corrected subgroup means in Table 4. Strictly speaking, the correlations from which the regression weights and multiple correlations are derived should be based on a single norm group. Since this was not possible, it was necessary to use the correlations in Table 2, based on the experimental group.

Even though it was not necessary to use the individual scores to calculate the corrected subgroup means for the experimental group, corrected PMA scores may be computed for any given individual by means of Table 5[3]. Assume that a person's corrected (P_c) score

[3] The derivation of this table is indicated in Section C of the material deposited with the American Documentation Institute. See Footnote 2.

TABLE 5
WEIGHTS USED IN CALCULATING INDIVIDUAL CORRECTED PMA SCORES

	V	S	R	N	W	I.Q.
W_P	1.489	1.054	1.555	1.134	1.200	1.514
W_S	.850	.011	.770	.346	.358	.752
W_M	.019	.197	.454	.136	.515	.424
W_R	.469	.194	.289	.204	−.293	.263
F	42.5	17.4	47.9	27.6	19.0	46.3

is desired for any given PMA score. Let his *T* scores on speed, memory, and motor rigidity be respectively C_S, C_M, and C_R. His P_c score is given by

$$P_c = PW_P - C_SW_S - C_MW_M - C_RW_R + F.$$

The values for the W's and for F are taken from the appropriate column of Table 5. For example, suppose a person's P_c score for the verbal factor is desired and that his verbal score is 47, and his speed, motor, and rigidity scores are, respectively, 53, 49, and 45. Using the values in the first column of Table 5 we have

$$P_c = 47(1.489) - 53(.850) - 49(.019) - 45(.469) + 42.5 = 45.4.$$

Curves for the corrected subgroup means are shown in Figure 3. When corrections are made

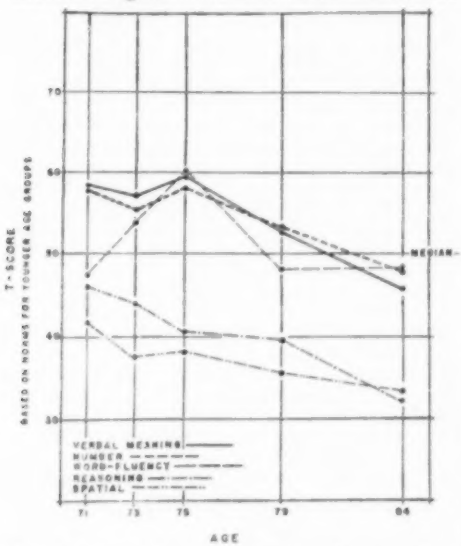

FIG. 3. MEAN *T* SCORES ON PRIMARY MENTAL ABILITIES CORRECTED FOR LOSS OF SPEED, MEMORY, AND FOR INCREASED MOTOR RIGIDITY

TABLE 6
STANDARD SCORES ON THE PRIMARY MENTAL ABILITIES TEST FOR SUBJECTS OVER EIGHTY YEARS OF AGE

Age	Verbal-meaning	Spatial	Reasoning	Number	Word-fluency
Females:					
80	58	33	39	39	40
80	56	32	48	56	49
80	56	39	42	49	38
81	42	36	30	50	58
82	54	35	33	58	39
82	34	27	27	27	40
83	50	27	27	52	41
88	35	37	34	35	67
Males:					
80	46	33	32	49	52
81	49	37	36	56	47
82	51	40	33	41	52
82	51	38	33	60	37
83	50	41	35	52	65
84	50	34	30	50	42
85	50	27	44	61	40
88	42	33	27	35	54

for losses in speed and memory and for an increase in motor-rigidity, the age curves for the Primary Mental Abilities flatten out appreciably. The Verbal, Number, and Word-fluency scores remain approximately 1 *SD* above the mean of the younger age groups at age 84. Scores on the Spatial factor, while significantly below the mean for the younger ages, show no appreciable change over the age range, 73–84.

The range of differences in Primary Mental Abilities scores for individuals of advanced age is of interest. Table 6 shows the uncorrected standard scores (with a mean of 50 and an *SD* of 10) on this test for the oldest subjects in the present group. Of the total of 16 in this age range, the verbal meaning scores of ten *S*s and the numerical ability scores of nine remain at or above the mean for an unselected population of 17-year-olds.

DISCUSSION

The retention of verbal ability shown here is consistent with the results of other studies. Owen's data (3) and those of Bailey and Oden, (2), both obtained by longitudinal methods, indicate that some verbal abilities continue to increase up to age 50. Doppelt and Wallace (4) have recently reported normative data for the Wechsler Adult Intelligence Scale on a fairly large and representative sample of older persons. The Vocabulary, Arithmetic, and Information tests show the smallest decline up

to age 70. For their oldest age group, with a mid-point age of 79.5, the Vocabulary score is only approximately .5 *SD*s below the mean of the reference population.

The studies most directly comparable with the present one are those of Schaie, Rosenthal, and Perlman (6), and of Sward (11). In the Schaie study, the oldest age group ranged from 71–78 years and the general educational level was considerably below that of the present group. On the *PMA* test, Schaie's older *S*s fell below the mean of the 17-year-old reference population on all factors. The lowest scores were obtained on the Space and Reasoning factors, as in the present study, and the highest scores were on Number and Verbal-meaning, as in the present case. However, the rank order of magnitude of these scores was different in the two studies. The earlier study showed the greatest loss on the Reasoning factor and the least on Number, while the present study showed the greatest loss on the Spatial factor and the least on Verbal-meaning.

Sward's study is of particular interest because his *S*s were drawn from university faculties and thus were fairly closely comparable to those used in the present study. His older group ranged in age from 60–79, with a mean CA of 66.2; his younger group ranged from 21–42, with a mean CA of 31.4. Unfortunately, from the standpoint of any direct comparison of his results with ours, Sward used a special test battery, the correlations of which with PMA scores are unknown. However, on all except two of his tests the older group did significantly less well than the younger. On his "word-meaning" test, the older group did as well as the younger. On a "synonym-antonym" test, the older *S*s were significantly better and on an "index of verbalization" which Sward constructed, the older *S*s also scored higher. These results are consistent with those obtained in the present study and lend support to the conclusion that, at least in a population of superior intelligence, verbal abilities remain at a high level into the middle seventies.

SUMMARY

Scores on primary mental abilities, speed, memory, and motor-rigidity were obtained for a group of fifty college graduates, ranging in age from 70–84 years. Differential decline in these abilities with advancing age is shown, with the earliest and largest losses occurring in memory, in speed, and in reasoning and spatial abilities. Word-fluency, verbal-meaning, and numerical abilities in this group of superior individuals remain well above the mean for young adults until the middle seventies. Verbal-meaning and numerical ability scores remain at or above the mean for the 17-year-old reference population in a majority of the *S*s over eighty years of age. Since this study is cross-sectional, rather than longitudinal, the possibility that selective factors may account for the high scores of the older *S*s cannot be ruled out.

REFERENCES

1. BAYLEY, NANCY. On the growth of intelligence. *Amer. Psychol.*, 1955, **10,** 805–818.
2. BAYLEY, NANCY, & ODEN, M. H. The maintenance of intellectual ability in gifted adults. *J. Gerontol.*, 1955, **10,** 91–107.
3. DENNIS, W. *Age and behavior.* Randolph Field, Texas: U. S. School of Aviation Medicine, 1953.
4. DOPPELT, J. E., & WALLACE, W. L. Standardization of the Wechsler Adult Intelligence Scale for Older Persons. *J. abnorm. soc. Psychol.*, 1955, **51,** 312–330.
5. OWENS, W. A. Age and mental abilities: a longitudinal study. *Genet. Psychol. Monogr.*, 1953, **48,** 3–54.
6. SCHAIE, K. W., ROSENTHAL, F., & PERLMAN, R. M. Differential mental deterioration of factorially "pure" functions in later maturity. *J. Gerontol.*, 1953, **8,** 191–196.
7. SCHAIE, K. W. The standard score method of computing intercorrelation matrices on the IBM type 604 electronic calculating punch. *IBM Technical Newsletter,* New York, 1955, No. 12.
8. SCHAIE, K. W. *Examiner manual for the test of behavioral rigidity.* American Documentation Institute, Library of Congress, Washington, D. C., Document No. 4651.
9. SCHAIE, K. W. A test of behavioral rigidity. *J. abnorm. soc. Psychol.*, 1955, **51,** 604–610.
10. SHOCK, N. W. Gerontology. *Ann. Rev. Psychol.*, 1951, **2,** 353–370.
11. SWARD, K. Age and mental ability in superior men. *Am. J. Psychol.*, 1945, **58,** 443–479.
12. THURSTONE, L. L., & THURSTONE, T. G. *Primary Mental Abilities Test.* Chicago: Science Research Associates, 1947.
13. WECHSLER, D. A standardized memory scale for clinical use. *J. Psychol.*, 1945, **19,** 87–95.

Received August 13, 1956.

JUDGMENTS OF CONSENSUS DURING GROUP DISCUSSION[1]

PHIL STONE AND JOE KAMIYA
University of Chicago

IF THE members of a group are asked to record privately their agreement or disagreement with a given proposition and to estimate the proportion of the group that shares the same opinion, the resulting data can be analyzed to reveal the degree of accuracy in the members' appraisal of their group's consensus. In several investigations (1, 2, 3, 4, 5) this technique, or some variant of it, has been used on a wide variety of opinion and attitude areas. Such studies have generally shown that individuals can estimate the consensus of their group with some degree of accuracy. Presumably, the estimates are based at least partly on knowledge acquired from past contact with whatever cues to the group's opinions that were available. The present study, with the general aim of shedding light on the manner in which the cues influence the estimates, undertakes a descriptive preliminary survey of judgments of group opinion in small discussion groups, where cues for the estimation of group opinion are provided through a discussion of the issues by the members themselves.

The questions explored are: Is the degree of estimation accuracy achieved through group discussion markedly different from that found in previous investigations where conditions of group discussion were not used? What is the nature of the errors? How are changes in estimates and in opinions related to the discussion content? Is accuracy increased by discussion?

METHOD

Subjects

The subjects (*Ss*) were six groups of high school students in Evanston, Illinois. The *N* of each group varied from 14 to 16. Both sexes were approximately equally represented in all classes. Ninety-two per cent of the students were either 15 or 16 years old, the remainder being 17.

Topics of Discussion

In order to establish the discussion itself as the primary source of cues for the estimates of group consensus, the topics presented were novel to the group. Pretests of comparable groups showed that the topics had probably not been discussed by the *Ss* before the study, but were of sufficient interest to maintain active discussion. In their order of presentation, the topics were: (*a*) whether the state law should be changed so that some high school juniors and seniors could drive the school buses at their school, (*b*) whether bicycle riders should be required by law to walk their bicycles across all intersections, and (*c*) whether women (adults) would be suitable as school bus drivers.

Procedure

The *Ss* were seated facing each other around a circle. Response forms were provided each *S*. The experimenter (*E*) began the session for each group by saying that the purpose of the study was to survey their opinions on some topics through group discussion. The *Ss* were told to try to reach a group decision on the issues presented them.

The *E* explained that during the discussion he would occasionally interrupt the discussion, and request each person to record on the form his opinion on the specific proposition then being discussed. Only "yes" or "no" opinions were permitted, corresponding to agreement or disagreement with the proposition. Immediately after indicating his own opinion, *S* was required to record his estimate of the number of persons in the group who held ***the same opinion as he did***. Each recording of an opinion and corresponding estimate is hereafter referred to as an "item."

The first item on each main topic was administered before any discussion occurred. The *S* was asked to give a favorable or unfavorable (yes or no) opinion and a consensus estimate on the topic as worded above. Subsequent items dealt with various facets of the main topic, as brought out by the discussants themselves. These items were careful formulations by *E* of the *S*'s contributions into concrete propositions. A sample item from the first topic: "High school junior and senior bus drivers wouldn't be able to handle the rowdies on the bus." "Answer 'yes' if you agree; 'no', if you disagree, and estimate the number of people who marked the form the way you did."

The extent of discussion allowed between items was varied. At times only one *S*'s contribution intervened; at the other extreme, as many as six different *Ss* were allowed to contribute before the next item was taken. Among all the items taken during the discussion of a main topic, some pairs of items were directly comparable, as a result of an *S* repeating a proposition which had been used as a previous item. These pairs were used to assess changes in the opinions and the estimates.

The *Ss* were informed in advance how long the discussion on each topic would last. At the end of the allotted period, *E* gave a final summary item and directed the discussion to the next main topic. The total time of discussion was about 45 minutes. The total

[1] This study is part of a project supported by a grant from the Social Science Research Council.

number of items taken in each group over all three major topics ranged from 18 to 26.

RESULTS AND DISCUSSION

Accuracy of Estimation

Presented first are measures of estimation accuracy over all items without regard to the nature of the discussion preceding the items. Two different measures were used. One measure was the difference between the estimated group opinion and the actual group opinion. The other measure was the Pearson product-moment correlation between estimated and actual opinion. These two measures were used in the analysis of the estimates of individuals as well as the pooled median estimates of the group.

Accuracy of individuals. For each *S*, the absolute difference between the estimated and actual group opinion was obtained on every item taken during the discussion period. The average of these differences over all items for a single *S* represented his average error score, in number of persons. The lowest average error obtained was 1.42 persons; the highest was 4.72 persons. The mean and *SD* of these average individual error scores are presented for each group in Table 1, Items 1 and 2.

For the correlational accuracy score of individuals over all items, *r*'s were computed between estimated and actual group opinion for each of 18 *S*s randomly selected from the six groups. The *r*'s ranged from .11 to .75 with a mean of about .40, values roughly comparable to the mean *r*'s of .69 and .44 reported by Travers (4) and Crow (2), respectively.

Accuracy of median estimates. For analysis of the estimates of the group as a whole, the median estimate was determined for each item. Since the task for each *S* was to estimate the number of *S*s who shared *his* opinion, a transformation was made before the median was computed to render the estimates of *all S*s into comparable estimates of the number of "yeses."

To measure group accuracy in terms of difference scores the absolute difference between the median estimate and the actual group opinion was determined for each item. The mean and *SD* of these error scores for each group are shown in Items 3 and 4 of Table 1. The mean error, expressed as a percentage of group size, is shown in Item 5. These values are also comparable to results found in other studies (3, 4).

As would be expected from the effects of pooling, the mean error of the estimates of the entire group is considerably smaller than the mean of the average error scores of single individuals shown in Item 1. However, further analysis indicates that the relative superiority of the group over the individual estimate is even greater than might be expected on these grounds: of the 88 *S*s tested, there are none with average error scores that were smaller than the mean error of the pooled estimates of their group.

In every group the correlation across all items between the median estimates of the entire group and the actual consensus was

TABLE 1
SUMMARY OF RESULTS

Item	Group					
	I	II	III	IV	V	VI
Number of Subjects	14	15	14	14	16	15
Number of Items	19	18	23	26	19	20
1. Mean of individual average error scores	2.43	2.65	2.57	2.44	2.57	2.90
2. *SD* of individual average error scores	.41	.68	.56	.53	.80	.79
3. Mean error of group median	1.26	1.33	1.61	.96	1.42	2.10
4. *SD* of mean error	.84	1.00	1.00	.65	.90	1.56
5. Item 3 as percentage of group size	9%	9%	11%	7%	9%	14%
6. Correlation between group estimate and actual consensus	.95	.96	.90	.95	.97	.90
7. Mean algebraic error of factions	+2.12	+2.47	+2.21	+2.52	+2.89	+2.42
8. *SD* of algebraic error	2.66	3.43	3.13	2.96	4.27	3.80
9. Mean absolute error of factions	2.68	3.42	3.05	3.16	3.76	3.65
10. *SD* of absolute error	2.11	2.86	2.37	2.28	3.56	2.42
11. Correlation between group estimate and actual consensus, own opinion held constant	.67	.56	.64	.72	.57	.63

.90 or higher (see Item 6, Table 1). These *r*'s are similar to the *r*'s of .82 and .93 found by Campbell (1) and Travers (4).

Accuracy and assumed similarity. Previous studies (1, 4, 5) have shown that *S*s' estimates of consensus are correlated with their own opinions. That is, those *S*s who share a given opinion on an issue give a larger estimate of the numbers holding their opinion than do those who hold the opposite opinion. To determine the extent to which this tendency for *S*s to assume similarity between themselves and others occurs under conditions of open discussion as created in this study, the median estimates of those whose own opinions were "yes" were computed separately from the median estimates of those whose opinions were "no." It was found that on every one of the 125 items in the entire study, the median estimate of one side was larger than the complement of the median estimate of the other side. The results of previous studies are thus clearly supported here.

In terms of accuracy, the foregoing finding means there is a general tendency for *S*s on each side of an issue to overestimate the numbers who share their opinion. The average net tendency toward overestimation per item is shown in Item 7 of Table 1. These figures are averages of the *algebraic* differences between the median estimate of each side and the actual number of *S*s on that side. A comparison of these differences with the absolute differences between the same measures (Item 9) clearly shows that the predominant error in each group is that of overestimation.

Consideration of the task confronting the *S*s makes it clear that the direction and extent of error in estimation can not be uniform for all sizes of "factions" (those who share one opinion). If all group members hold the same opinion, the only error of estimation that can possibly occur is underestimation, since the highest estimate that can be given is limited to the total number of persons in the group. At the other extreme, if only one member holds a given opinion, he can err only in the direction of overestimation.

If such constraining effects of group size were the only factors influencing the errors in the estimates, equal but opposite errors in the median estimates of the two extremes of faction size would result. However, as could be predicted from the observed widespread tendency to overestimate, such symmetry in the errors did not occur in the data. Analysis showed that factions comprised of about 10% of the group gave median estimates that averaged about 55-60% of the group size, while factions comprised of 90% of the group gave median estimates that were about 80% of the group size. The estimates of the intermediate sized factions fell approximately on a straight line between these two extremes, with greatest accuracy of estimation being found in factions comprised of approximately 75% of the group. Hence, it is clear that all but the largest factions tended to overestimate their size.

The fact that *all* factions, regardless of size, assumed on the average that more than half the group shared their own opinion (or that less than half of the group held the opposite opinion) accounts in part for the high degree of accuracy of the median estimate of the entire group, since a majority of the estimates on which the group median is based comes from the majority side, and the pooled estimate of the entire group is therefore partly determined by the actual consensus. The accuracy of the pooled median estimate is thus spuriously high when taken as an index of the judgmental achievement of the group. As an empirical fact, however, the high accuracy of median estimates should lend itself to the use of estimates for the indirect measurement of attitudes in small groups.

An attempt was made to determine whether any correlational accuracy would remain if the factor of assumed similarity were excluded. The *S*'s own opinion was held constant by computing the accuracy correlations of the "yes" factions and the "no" factions separately across all items, and then averaging the two correlations. The results, shown in Item 11 of the table, indicate appreciable judgmental accuracy apart from the spurious effects of assumed similarity, with *r* ranging from .56 to .67 among the six groups.

The Influence of Group Discussion on Estimates and Opinions

The results to be reported focus on the relationship between the views expressed during the course of the discussion and the accompanying changes that occurred in the

estimates and opinions. All six groups were combined in this analysis. The changes were assessed by comparing the estimates and opinions taken on an item before discussion with those taken on a comparable item after discussion. From the six groups, a total of 49 comparable before-and-after item pairs was found.

Overall favorableness or unfavorableness toward the propositions presented of the views expressed in discussion was rated on a seven-point scale by two judges using verbatim transcripts. All relevant expressions of opinion intervening between the comparable items were rated as a single unit. The judges based their ratings on the number of different *S*s speaking as well as the content of each *S*'s speech. The inter-judge correlation over the 49 discussion units was .82. The average rating of the two judges on each item was transformed into a percentage according to its relative location within the range of all the ratings, in order to provide an index of the favorability of discussion that could be compared directly with the degree of favorableness of the group opinion measures.

Changes in estimates. Since the views expressed in support of a proposition (or in opposition to it) during discussion provide cues for the estimation of consensus, it is reasonable to expect the median estimate of the group to change in the direction of the expressed support, unless the expressed support is representative of the initial estimate. The expressed support was considered representative if it did not differ from the initial estimate by more than 10% of group size. Eliminating these representative cases, the median estimate shifted toward the level of expressed support in 35 out of 41 cases ($p < .001$.)[2]

Changes in actual opinion. Since the views expressed during discussion may affect opinions as well as estimates, a parallel analysis was undertaken of the changes in the opinions. The results were again clear cut. In 32 out of 41 cases, the actual group opinion shifted in the direction of the expressed support. This obtained proportion, evaluated against its chance level of .53 (obtained by the same procedure as for the estimates) is significant at the .001 level.

Other findings. The correlation between changes in the median estimates of the group and the changes in actual group consensus was found to be .89, significant at the .001 level.

It might be expected that the estimates would shift more in the direction of the expressed support than would the opinions, on the ground that the *S*s are likely to think others are more affected by what is said than they are themselves. However, the results from several methods of analysis point to no significant difference between the change scores for the two measures.

No significant tendency of error reduction was found. In the 49 units of discussion, error (expressed as the sum of the members' absolute errors on an item) was reduced only 22 times. Moreover, no tendency was found for error reduction to be related to the number of people who spoke during the discussion unit.

Although the error is not reduced, the discussion might nevertheless tend to produce a convergence of estimates toward a common level. A tendency was in fact found for the variance among the individual's estimates to be reduced by discussion ($p = .08$).

Interpretation. That accuracy was not increased by the discussion is understandable in light of the high degree of similarity observed in the changes of the estimated and actual consensus. The similarity of the changes, in turn, is in part accountable for by the previously observed relationship between the *S*s' estimates and their own opinions. Given this relationship, *changes* in the *S*s' opinions should be accompanied by changes in their estimates.

Whatever the reasons may be, it is apparent that not only do the expressions of opinion influence the estimates; they influence the group consensus itself, and in substantially the same manner. The types of conditions upon which this dual effect depends needs to

[2] For assessing the significance of this proportion, the value for the chance proportion under the null hypothesis was determined empirically, following the method used by Crow (2) for a somewhat different problem. The relative frequency with which the obtained changes occur in the expected direction was determined, using level of expressed support in every *other* discussion unit in the study as a base of prediction. This gave a total of 49×48, or 2352, comparisons. The proportion of cases in which the estimates shifted in the direction of these other ratings of discussion units was .57. This value was taken as p in the binomial distribution, against which the obtained proportion of 35/41 was evaluated.

be explored. The findings of the present study suggest that rather complex interactions may be involved among expressions of opinion, group consensus, and judgments of consensus, and that the common conception of group discussion as a means of increasing the understanding of the group members' views needs qualification.

SUMMARY

Six groups of high school students held discussions during which the individual's opinions and estimates of the group consensus were privately taken on the issues considered.

The degree of estimation accuracy achieved was not noticeably different from the accuracy found by other investigators under conditions of no discussion. Tendencies among *S*s to overestimate the number of group members who share their opinion similar to ones previously reported, were also found. Even minorities comprised of only about 10% of the group tended to estimate that more than half the group shared their opinion.

The views expressed by the *S*s during discussion influenced both the median estimate of consensus and the consensus itself. The influence on both was highly similar in that (*a*) the amount of change in the actual consensus was correlated with the amount of change in the estimated consensus, (*b*) both the estimated and actual consensus tended to change in the direction of the views expressed, and (*c*) the amount of change in the estimated consensus was not significantly greater than the amount of change in the actual consensus.

Over the course of discussion, the estimates did not tend to become more accurate. There was a slight tendency toward reduction in the variance of estimates among the *S*s.

REFERENCES

1. CAMPBELL, D. T. *A study of leadership among submarine officers.* Columbus: Ohio State Univer. Res. Found., 1953.
2. CROW, W. J. A methodological study of social perceptiveness. Unpublished doctor's dissertation, Univer. of Colorado, 1954.
3. NEWCOMB, T. *Personality and social change.* New York: Dryden Press, 1943.
4. TRAVERS, R. M. W. A study in judging the opinions of groups. *Arch. of Psychol.*, 1941, No. 266.
5. WALLEN, R. Individuals' estimates of group opinion. *J. soc. Psychol.*, 1943, **17**, 269–274.

Received August 16, 1956.

COGNITIVE PROCESSES AND INTERPERSONAL PREDICTIONS[1]

HOWARD LEVENTHAL[2]

University of North Carolina

THE predictions of an observer about the behavior of another person are a product of the observer's hypotheses or conceptualization regarding the particular individual. The accuracy of his predictions depends upon the validity of the conceptualization. His hypotheses are, in turn, a function of the information available about the person to be judged, and of the judge's typical way of categorizing information or forming concepts about others. These assumptions were developed into three hypotheses which the present study was designed to test.

A more accurate impression or conceptualization of another person should be possible when more information is given the observer about the individual, and a more accurate prediction of his behavior should therefore result. *Hypothesis 1. The more information provided judges about a person, the more accurate are their predictions of his behavior.*

The judge's typical concepts or ways of grouping others determine the kinds of discriminations and consequent predictions he makes about them. Judges with a larger number of concepts for grouping others (hereafter called Complex Judges) should develop more complete and more accurate impressions of an individual than those who use few concepts (hereafter called Simple Judges). *Hypothesis 2. Complex Judges more accurately predict behavior than Simple Judges.*

Since one function of information is to provide the setting for fine or multifactor discriminations, one would predict that the impressions developed by Complex Judges would be more complete and representative of a particular person when they are given additional information. Simple Judges, however, should continue to classify data about the person observed into the few categories available to them and their impressions should not increase in accuracy with amount of information provided. *Hypothesis 3. With increasing information, Complex Judges show a relatively greater increase in predictive accuracy than Simple Judges.*

METHOD

The design required a pool of subjects to serve as judges and interviewees, who varied in respect to simplicity or complexity of interpersonal concepts, and in the amount of information received. Judges were given information about another person by means of tape-recorded interviews, and then they predicted the other person's responses to a questionnaire.

Selection of Judges and Interviewees

Bieri's modification of the Kelly Role Repertory Test (1, 3) was used to rate subjects for conceptual simplicity or complexity. Subjects (*Ss*) were given a list of 12 different categories of people (e.g., a person you'd like to know better; the most ethical person you know, etc.). The *S* named an acquaintance or relative to whom each description could be applied. He compared selected triads of these 12 individuals, and for each triad he was instructed to provide a personal characteristic or trait in regard to which two of the persons were similar to one another but different from the third. Any trait or psychological concept could be used for the differentiation. The *S* then checked the two like individuals. His next task was to see if the concept applied to any of the other nine individuals not in the original comparison, and to check any to whom the concept applied. The *Ss* made twelve comparisons, using a different specified triad for each.

Scorable protocols on this test were obtained from 253 male students in elementary psychology, and scored for overlap or simplicity in use of concepts. To the extent that the subject checked the same individuals as examples of the constructs he had formed in the successive comparisons of triads, he obtained a high score on simplicity. Scores obtained ranged from 0 to 82, with a distribution skewed toward high, or simple scores. Subjects scoring 20 or above were classified as simple, and those scoring between two and six as complex. From these extreme groups, 14 *Ss* were selected as interviewees (7 simple and 7 complex), and 56 as judges (28 simple and 28 complex).

To determine whether simplicity-complexity was related to intelligence, the A.C.E. scores for a randomly selected sample of these *Ss*, 95 in all, were correlated

[1] This study is a condensed version of a dissertation submitted to the University of North Carolina in partial fulfillment of the requirements for the Ph.D. degree in June of 1956. Thanks are extended to Dr. June E. Chance of the Department of Psychology, and to Dr. David Cox of the Department of Bio-Statistics for their invaluable help and criticisms.

[2] Now with the Behavioral Studies Section, U.S.P.H.S.

with their simplicity-complexity scores. The obtained correlation ($r = -.04$) was not significantly different from zero.

Information Provided

The information from which judges were to make their predictions was obtained by tape recording interviews with seven simple and seven complex individuals. Four content areas were covered in the interview: (*a*) family background—description of parents and siblings; (*b*) school and job experiences—descriptions of teachers and bosses as well as relevant activities; (*c*) values—description of girl friend, and ideal mate, future plans; (*d*) self-description—how others see self, desires for personal change, and attitudes toward others. The information from topics *a* and *b* above, about 15 minutes of the tape, constituted partial information, while the entire interview of about 35 minutes constituted the condition of complete information.

The Prediction Questionnaire

Upon the close of the interview, the interviewee gave his answers to a 38-item questionnaire which then served as the criterion for the accuracy of the judges' predictions.

This questionnaire consisted of 38 four-choice items concerning behavior in social, work, and school situations. The questionnaire was administered in the elementary psychology classes to determine the distribution of responses to the items under conditions of self-description. Completed questionnaires were obtained from 236 males in addition to the 56 who later served as judges.

The distribution of responses for the group of 236 showed definite stereotypy of self-description, indicating that the chance level of accuracy would not be the same for all item alternatives. In order to compensate for the advantage this stereotypy would otherwise have given a judge who relied on group norms rather than on the interview information for his predictions, a weighting formula was developed which gave greater weight to accurate predictions the more the interviewee's response deviated from the group stereotype and the more the judge's prediction deviated from the use of "normative concepts." The proportion of times judges predicted each alternative was tabulated for every item, and weights to be applied to correct predictions of an alternative were derived according to the following formula: one minus the proportion of self choices divided by the proportion of times the choice is predicted by judges. As the proportion of times an alternative is used for self-description or prediction increases, the weight decreases. The use of the formula was intended to give greater weight to predictions that reflected discrimination of qualities characteristic of the *individual* interviewee.

Design and Procedure

The experiment was designed so that both Simple and Complex Judges judged a simple and a complex interviewee, with partial and complete information, and with each of these combinations of variables in first and second position. The design (Table 1) called for four pairs of judges, each pair composed of a Simple and Complex Judge, and one pair of interviewees, a Simple and a Complex. This design was replicated 7 times with a total of 56 judges and 14 interviewees participating.

TABLE 1

EXPERIMENTAL DESIGN SHOWING TYPE OF INTERVIEWEE, AMOUNT OF INFORMATION, AND SEQUENCE OF JUDGING FOR SIMPLE AND COMPLEX JUDGE PAIRS

	Judge Pairs			
	1	2	3	4
Order Judged	Simple Complex	Simple Complex	Simple Complex	Simple Complex
First				
Amt. of Inf.	15 min.	35 min.	15 min.	35 min.
Interviewee	simple	simple	complex	complex
Second				
Amt. of Inf.	35 min.	15 min.	35 min.	15 min.
Interviewee	complex	complex	simple	simple

Two judges, one Simple and one Complex, were scheduled for each of the judging sessions. Before listening to the recorded interviews, judges studied the content of the questionnaire items. They were then instructed to listen to each of two interviews, and to formulate as accurate an impression of the interviewees as possible. After hearing each recording, the judges answered the questionnaire in the way they thought the interviewee had. An incentive award of two dollars was offered to the judge who was most accurate for one interviewee, and an eight dollar prize if he was most accurate for both.

RESULTS AND DISCUSSION

Hypothesis 1. Predictions made with complete information were more accurate than those made with partial information ($p < .05$), when unweighted scores are considered. This effect was not significant for the weighted accuracy scores. Perhaps complete information provided the judges with a clearer idea of the interviewee's subgroup membership, rather than of his idiosyncratic characteristics.

Hypothesis 2. There was no difference between Simple and Complex Judges in the total number of correct predictions. However, when compared on weighted accuracy, Complex Judges showed a tendency to be more accurate ($p < .10$). This trend is consistent with the interpretation that Complex Judges are more likely to base their predictions on the characteristics of a particular interviewee.

Hypothesis 3. A comparison of the changes in unweighted accuracy scores revealed that Simple Judges showed relatively greater improvement than Complex Judges when

TABLE 2
ANALYSIS OF VARIANCE, UNWEIGHTED ACCURACY SCORES

	df	Variance	*F*
Between Judge Source			
Between Subject Pairs	6	52.31	3.76***
Judge Type	1	.08	
Information × Order	1	2.01	
Information × Subject Type	1	6.51	
Order × Subject Type	1	16.51	
J. Type × Subj. Pairs	6	13.00	
J. Type × Inf. × Order	1	8.58	
J. Type × Inf. × S. Type	1	13.58	
J. Type × Order × S. Type	1	1.08	
Residual Between	36	13.91	
Within Judge Source			
Subject Type	1	40.08	
Subj. Type × Subj. Pairs	6	31.87	4.60***
Information	1	33.22	4.80**
Order	1	25.08	3.62*
J. Type × Inf.	1	19.72	2.85*
J. Type × Order	1	33.22	4.92**
J. Type × S. Type	1	.72	
Inf. × Order × S. Type	1	8.58	
Residual Within	43	6.92	
Residual Bet./Residual Within	36/43		2.01**

* $.10 > P > .05$.
** $.05 > P > .01$.
*** $.01 > P > .001$.

judging with complete as compared with partial information ($p < .10$). The increase in unweighted accuracy with complete information was, therefore, largely due to increased accuracy by Simple Judges.

These data raise two questions: (*a*) How did Simple and Complex Judges differ in the process by which they arrived at their predictions, and (*b*) What role did perception of individual interviewee characteristics play in prediction? How can we understand the trends for Complex Judges to predict more atypical items and for Simple Judges to improve more with complete information? Could these differences in accuracy have been a function of actual similarity between the judges and the interviewees in their own responses to the questionaire? Analysis of similarity in self-response indicated no significant differences in judge–interviewee similarity for any experimental condition.

Since Bieri (1) found greater "assimilative projection," i.e., prediction of similarity, in Simple Judges, it is possible that prediction of similarity can account for the accuracy of Simple Judges in the present experiment. An analysis showed significantly greater prediction of similarity between interviewee's and judge's responses by Simple Judges than by Complex Judges ($p < .05$). Therefore, the components of accuracy (the proportion of actual response similarities accurately predicted, and the proportion of actual response differences accurately predicted) were analyzed. The trends showed that Simple Judges correctly predicted a greater proportion of similarities and the Complex Judges correctly predicted a greater proportion of response differences. However, neither trend was significant.

Though Simple Judges predicted significantly greater similarity than did Complex Judges, Complex Judges were more able to order the interviewees in terms of similarity to themselves. Correlations of actual similarity with amount of similarity predicted by the judge were computed separately for each combination of experimental conditions and z scores averaged. The correlations were, $r = .49$

TABLE 3
ANALYSIS OF VARIANCE, WEIGHTED ACCURACY SCORES

	df	Variance	*F*
Between Judge Source			
Between Subject Pairs	6	260.34	2.66**
Judge Type	1	375.22	3.84*
Information × Order	1	82.05	
Inf. × Subj. Type	1	12.04	
Order × Subj. Type	1	30.93	
J. Type × Subj. Pairs	6	122.24	1.25
J. Type × Inf. × Order	1	85.16	
J. Type × Inf. × S. Type	1	4.58	
J. Type × Order × S. Type	1	40.87	
Residual Between	36	97.77	
Within Judge Source			
Subject Type	1	175.20	
Subj. Type × Subj. Pairs	6	406.87	4.62***
Information	1	188.45	2.14
Order	1	385.32	4.38**
J. Type × Information	1	18.14	
J. Type × Order	1	438.06	4.97**
J. Type × Subj. Type	1	7.22	
Inf. × Order × Subj. Type	1	6.75	
Residual Within	43	88.06	
Residual Bet./Residual Within	36/43		1.11

* $.10 > P > .05$.
** $.05 > P > .01$.
*** $.01 > P > .001$.

($p < .01$) for Complex Judges, and $r = .24$ (not significantly different from zero) for Simple Judges. These results support the possibility that Complex Judges predicted on the basis of an impression of the individual interviewee that was generated from the information provided him. Although this interpretation is in agreement with the finding that Complex Judges were more accurate on individually relevant items, the data are more suggestive than conclusive.

Observing judge performance as a function of the other independent variables may further clarify the nature of the prediction process as well as differences between the Judge types. While the increase in accuracy with full information was greater for Simple Judges than for Complex, both types of judges showed greater accuracy under this condition. The increase in accuracy cannot be attributed to predictions of similarity since predictions of similarity decreased significantly ($p < .05$) with complete information. Complete information also resulted in a decrease in the proportion of real differences incorrectly predicted as similarities ($p < .01$), and an increase in the proportion of correctly predicted differences ($p < .001$). These results may be a function of the content of the information as well as of the amount of information. Complete information alone disclosed the interviewees' self-perceptions to the judges. A possible reason for complete information being more helpful to Simple and Complex Judges is that the material on the interviewees' self-perceptions may have been very relevant for correcting the kind of errors the Simple Judges were making, i.e., not differentiating the interviewees from themselves. On the other hand, Complex Judges, when making predictions with complete information, may have over-differentiated between themselves and others. Over-differentiating the interviewees is likely to result in error, since people tend to respond similarly to the questionnaire. Trends in the data showed that with complete information, Complex Judges decreased their accurate predictions of similarity, and increased their inaccurate predictions of differences. Perhaps Complex Judges would have been more accurate if they knew more about the normative responses to the items, or if the criterion had recorded more individualistic responses.

TABLE 4
ANALYSIS OF VARIANCE, PREDICTED SIMILARITY BETWEEN JUDGE AND INTERVIEWEE

	df	Variance	*F*
Between Judge Source			
Between Subject Pairs	6	38.85	
Judge Type	1	150.89	4.14**
Information × Order	1	2.89	
Inf. × Subj. Type	1	48.89	
Order × Subj. Type	1	11.57	
J. Type × Subj. Pairs	6	28.27	
J. Type × Inf. × Order	1	14.29	
J. Type × Inf. × Subj. Type	1	137.29	3.77*
J. Type × Order × Subj. Type	1	22.32	
Residual Between	36	36.44	
Within Judge Source			
Subject Type	1	6.04	
Subj. Type × Subj. Pairs	6	23.08	1.43
Information	1	89.29	5.53**
Order	1	54.32	3.36*
J. Type × Inf.	1	26.03	
J. Type × Order	1	24.14	
J. Type × Subj. Type	1	20.57	
Inf. × Order × Subj. Type	1	36.57	2.26
Residual Within	43	16.15	
Residual Bet./Residual Within	36/43		2.26***

* $.10 > P > .05$
** $.05 > P > .01$.
*** $.01 > P > .001$.

Judging interviews in succession produced effects upon accuracy and predicted similarity. The trends for the second interview were toward decreased accuracy, ($p < .10$ for unweighted accuracy and $p < .05$ for weighted accuracy) and greater prediction of similarity ($p < .10$). The decrement in accuracy was largely confined to Complex Judges ($p < .05$ for the Judge Type X Order interaction on both accuracy measures). Simple Judges did about equally well with interviews judged in either position. The deterioration of prediction among Complex Judges may well reflect their inability to continue to respond to the individual interviewee's characteristics. If a Complex Judge tried in his predictions to distinguish the second interviewee from both himself and the first interviewee, the alternatives remaining for his choice were likely to be those least frequently selected for self-description.

In general, it appeared that Complex Judges responded in terms of differences between people, while Simple Judges responded in terms of similarities. Viewing people as similar to oneself as an initial hypothesis may help

Simple Judges to understand and adjust to their own interpersonal environments. They differentiate only when sufficient information is available. Complex Judges, on the other hand, seem to seek a unique characterization of a person and to be fairly able to pick out aspects of his individuality even with little information. It appears that the general tenor of the experimental instructions, which implied judging the uniqueness or differences between individuals, had a stronger effect upon the Complex Judges. The Simple Judges of the present study resemble Lundy's (5) *Ss* who were instructed to attend to their own reactions while evaluating another person, and Fiedler's (2) *Ss* who rated high in assumed similarity between opposites.

Do we have data that can tell us whether the judges actually perceived and responded in terms of characteristics of the individual interviewees, as well as responding in terms of their expectations of how college males would answer a questionnaire? The changes in accuracy and predicted similarity with information give support to the view that information as well as expectations determined the predictions. Had data on the effects of information and order not been obtained, our conclusions might have been different. We would have seen that accuracy correlated with similarity, and that variables such as the proportion of similarities and proportion of differences correctly predicted did not correlate, and might have concluded that information specific to an individual interviewee was not significant for accurate prediction.

The results suggest that research directed toward explaining the prediction process should include (*a*) more specific treatment of judges' self-constructs or the ways in which they compare themselves to others, (*b*) manipulation of the similarity between judge and judged as well as control of the similarity of their responses to group norms, (*c*) more careful control of variations in the content of information given to judges, and (*d*) systematic variation of the relevance of the content of the information to the content of the predictions.

SUMMARY

A set of interrelated variables thought to influence predictive accuracy were studied experimentally. A model treating social perception as an example of concept formation suggested the combined manipulation of stimulus and individual difference variables. It was expected that judges possessing a number of differentiable concepts regarding others (Complex Judges) would be better predictors than judges having undifferentiated interpersonal constructs (Simple Judges). It was hypothesized that given more information about a person, judges would predict better and that Complex Judges would best utilize full information. Scores on the Kelly Role Repertory test defined the dimension of conceptual simplicity-complexity. Information was presented by means of recorded interviews, each type of judge judging each type of interviewee, and judgments were made on the basis of two different amounts of information. The results indicated that though Simple and Complex Judges were about equal in predictive accuracy, they differed in the ways in which this accuracy was achieved. Simple Judges tended to assume similarity between themselves and others, while Complex Judges tended to differentiate between themselves and others. The increase in accuracy resulting from the provision of more complete information also seemed to indicate that the judges responded in terms of the individual characteristics of the interviewees, as well as in terms of group norms.

REFERENCES

1. BIERI, J. Cognitive complexity-simplicity and predictive behavior. *J. abnorm. soc. Psychol.*, 1955, **51**, 263–268.
2. FIEDLER, F. E. Assumed similarity measures as predictors of team effectiveness. In A. P. Hare, E. F. Borgatta, and R. F. Bales (Eds.), *Small groups*. New York: Alfred A. Knopf, 1955, Pp. 252–265.
3. KELLY, G. A. *The psychology of personal constructs*. Vol. I. *A Theory of Personality*. New York: W. W. Norton, 1955.
4. LEVENTHAL, H. Effects of setting factors and judge's interpersonal constructs on prediction behavior. Unpublished doctor's dissertation, Univer. of North Carolina, 1956.
5. LUNDY, R. M. Assimilative projection and accuracy of prediction in interpersonal perceptions. *J. abnorm. soc. Psychol.*, 1956, **52**, 33–38.

Received September 19, 1956.

CONFORMING BEHAVIOR UNDER TWO CONDITIONS OF CONFLICT IN THE COGNITIVE FIELD[1]

FRANZ SAMELSON[2]

Mental Health Research Institute, University of Michigan

ONE of the oldest topics in experimental social psychology is the effect of a communication from other persons on an individual's opinion or judgment: suggestibility and conformity. While the early experiments simply demonstrated this effect in various situations, later ones have shown that the amount of conformity obtained is related to three clusters of factors: (*a*) the task stimulus, its difficulty and ambiguity (5, 14); (*b*) the social influence, its source, content, and discrepancy with either the true facts or the individual's initial judgment (2, 9); and (*c*) characteristics of the individual, permanent or temporarily induced (6, 12). While the phenomenon is well enough established—in fact, seems to be one of the cornerstones of social psychology—the interpretation of the change processes, their "rationality" or "irrationality," has been the subject of controversy (2, ch. 14).

In an attempt to clarify the issue, the distinction between two types of forces, viz., "normative" and "informational" social influences, has been suggested recently by Deutsch and Gerard (7), thus linking the topic with the theory of reference groups (11). The importance of *normative* social influences for conforming behavior is shown in the experiment by Deutsch and Gerard, and a similar one by Mouton *et al.* (16), in which anonymity of the subjects (*Ss*) reduced the amount of conformity obtained significantly, but by no means completely.

Informational social influences, or the creation of a frame of reference, as in Sherif's classical study with the autokinetic phenomenon (21), seemed to be especially important in the realm of "social reality," i.e., in regard to issues which are not subject to direct observation or independent test. Yet Asch's famous experiment on conformity (2, ch. 16) and its subsequent modifications (*v.s.*) seem to indicate that even in simple perceptual tasks (including immediate memory) wrong information, contained in other *Ss*' responses, had some power to change the judgments of the naive observer. The implication seems to be that even simple perceptions may need some "consensual validation." This point touches on the old philosophical problem of "reality" and knowledge about reality, and the suggestion given is quite similar to the philosophical position[3] that sense experience is transformed into experience of reality "out there" of material objects only after its corroboration by other observers, or after repeated independent observations by the individual, especially in the form of what Allport (1) has called denotational procedures. (Typically, this latter course of action is made impossible for the *Ss* in conformity and suggestion experiments—a restriction which may seriously limit the generality of the results.) In other words, viewing perceptions as hypotheses (4), one may expect that even after their initial formulation they may still be susceptible to confirmation or infirmation from reports of other observers.

If these premises are accepted and the discussion restricted to the informational component of the influence attempt, the following analysis of Asch's experiment seems plausible. The task stimuli (lines to be compared) were by themselves quite free from ambiguity. Yet this clearness of the perception combined with the similarly unequivocal but contradictory evaluation by the other observers to create an unstable total situation for the naive *S*. He was, in fact, required to choose between two mutually exclusive messages, yet had no independent basis to discount

[1] This paper is based on a doctoral dissertation submitted to the Psychology Department of the University of Michigan. The writer wishes to express his gratitude to Daniel Katz, chairman of the doctoral committee, and its members, especially Theodore M. Newcomb and Russell A. Clark. The study was carried out under a research training fellowship granted by the Social Science Research Council.

[2] Now at Kansas State College.

[3] For a recent discussion, see the article by Fritz (8) and its distinction between S-statements (statements about sense-data) and M-statements (about material objects).

either of them. Since at the outset the task seemed quite simple, the initial expectation of observer agreement was presumably very high. The surprising discrepancy between the *S*'s perceptions and those of a number of other *S*s, unanimous among themselves, created a conflict in the cognitive field which could not be resolved in a satisfactory manner by simply denying or disregarding the socially transmitted information. Instead, *S*s tried to *account* for the disagreement by means of additional hypotheses which, because of their *ad hoc* nature, left the *S*s rather disturbed and dissatisfied whether they had yielded or remained independent.

On the basis of these considerations, the following hypothesis was proposed: The amount of conformity in situations of the type studied by Asch is reduced if *S*s are given secondary information which, if utilized, reconciles the contradiction between their own perception of the task stimuli and the majority judgments, and thus enables them to account for the disagreement. This hypothesis is somewhat similar to the (untested) Hypothesis VI of Deutsch and Gerard: "The more uncertain the individual is about the correctness of the judgment of others, the less likely he is to be susceptible to informational social influence in making his judgment" (7, p. 630).

METHOD[4]

In order to test this hypothesis, two conditions were designed under which *S*s found themselves in disagreement with a secretly instructed majority. Their task was to identify simple visual stimuli that were presented tachistoscopically but above threshold duration. This method of presentation, rather than Asch's original procedure, was used since it made the intended manipulations more plausible[5]. In the control, or Full Conflict (F), condition, *S*s were put—as were Asch's *S*s—into the conflict situation without any additional information. In the Reduced Conflict (R) condition[6], the *S*s were first given the opportunity to infer that their recognition thresholds were lower than those of the other participants. Subsequently, they found themselves in the same conflict with the majority as the control *S*s.

Subjects

Fifty beginning psychology students, all male freshmen and sophomores at the University of Michigan, were used as *S*s. They were assigned at random to either of the two conditions. This prior assignment, together with scheduling difficulties, eventually left the two groups slightly imbalanced ($N_F = 26$, $N_R = 24$).

Task

Each of these *S*s met with four other *S*s, who were in fact secretly instructed accomplices of the *E*, for an experiment "on the span of apprehension." Their task was to recognize a four-consonant nonsense syllable presented at .2 sec., and to select it from four alternative syllables presented approximately 1 sec. afterwards, and left on the screen till all *S*s had responded. The three distractors had been constructed in such a way that one letter and/or its position was changed. For example, for the stimulus syllable FDMR, the wrong alternatives were: PDMR, PDRM, FDRM. All five participants announced their choice publicly in a fixed sequence by calling out the number of the alternative they believed to be correct. The naive *S* was always fourth to report. Twenty trials with different syllables were given. On twelve, all four accomplices gave the same wrong answer. On the remaining eight trials, spaced in an irregular order over the experiment, they chose unanimously the correct alternative. The sequence of critical and neutral trials is indicated in Figure 1.

Manipulation of Conditions

The experimental treatment, designed to provide *S*s in the R condition with information potentially reducing the conflict between the physical and the social stimuli, consisted of the following procedure. Prior to the test trials, the groups in the Reduced Conflict condition were given simulated practice trials. Each of three slides with a five-digit number was shown to the group repeatedly at decreasing speeds, beginning with .01 sec. All *S*s were asked to report, out of sequence, what they had been able to recognize. The accomplices, however, had been instructed to act as though they were unable to recognize anything at the initial fast speed, and to give correct responses only *after* the naive *S* had given an accurate report, and after the speed of presentation of the *same* slide had been decreased twice. Through this procedure, the *S*s in the Reduced Conflict condition presumably gained the impression that their recognition threshold was lower than those of the other

[4] For a more detailed description of procedures and results, see (19).

[5] Similar modifications of Asch's procedures have been used effectively by others, e.g., Olmstead and Blake (18).

[6] These titles of conditions refer to the predicted amounts of vacillation between independent and conforming responses. Continued independent behavior, as predicted for the Reduced Conflict group, may well arouse considerable anxiety, so that an over-all measure of conflict experienced by the *S*s would not necessarily show differences between conditions corresponding to these titles.

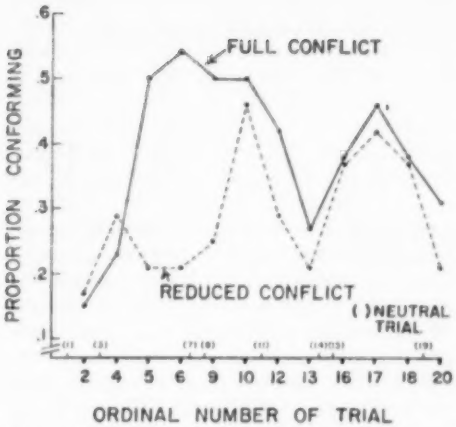

FIG. 1. PROPORTION OF CONFORMING RESPONSES ON SUCCESSIVE CRITICAL TRIALS IN TWO EXPERIMENTAL CONDITIONS

observers. In the Full Conflict condition, no such practice was given. The objective informational as well as normative social influence toward conformity, i.e., the majority's agreement on wrong responses in the test trials, remained identical over both conditions.

Apparatus and Setting

A three-hundred watt projector with a Keystone flashmeter attachment was used for presentation of the stimuli. These consisted of four capital letters typed on cellophane and mounted for projection. The distance from screen to projector was 24′, to the *S*s about 19′. The *S*s sat in a row, separated from each other by small screens in order to reduce the demands on the accomplices' acting skill. The back half of the room was illuminated in order to permit *S*s to write during the session, as well as to avoid excessive contrast in illumination level.

Additional Data

After each judgment, *S*s indicated their uncertainty about the correct response on a five-point rating scale from "absolutely certain" to "probably wrong." One of the objectives of this measure was to check, by an intra-individual comparison, whether *S*s would differentiate in their ratings between neutral trials and trials on which they had conformed. Since outwardly the response pattern of the group would be identical in both cases, higher uncertainty ratings on the conformity trials would be expected if the *S*s had some awareness of conforming to *wrong* majority responses. Such a result would give support to the assumption that *S*s did not conform simply because they could not recognize the visual stimuli.

Immediately after the last trial each *S* was given a questionnaire containing three rating scales and two open-ended questions. The first two ratings dealt with the difficulty of the task and the *S*'s visual acuity. On the third scale he was asked to indicate a judgment of his ability for a task of the type just performed. This question was designed to test whether the experimental manipulations had been perceived in the intended way. The remaining two items were a question asking whether the *S* had been influenced by someone else's responses, and a conditional question about his explanation of the disagreement with other *S*s "in case such disagreements had occurred."

This task was followed by oral questioning of the naive *S* under the pretext of a group discussion of the experiment. The topics covered in this interview were the *S*'s awareness of the disagreement with the majority, his feelings about who had been right, and his way of accounting for the discrepancy, gradually leading to the question whether he had developed any suspicions about the behavior of the majority. None of the *S*s had been informed beforehand by rumor, and none had surmised the extent of the deception during the experiment, although a few had begun to suspect that the experiment might not be quite what it purported to be. Finally, the purpose and procedures were fully explained to every *S* since all of them showed some signs of disturbance about the puzzling course of events.

Personality Data

Some weeks earlier, in the context of another experiment, measures of the *S*s' Achievement and Affiliation motives had been obtained. These projective motivation measures are based on stories written in response to TAT pictures and scored according to conventions developed by McClelland *et al.* for need Achievement (15), and by Atkinson *et al.* for need Affiliation (3). Need Achievement scores are presumably indices of the strength of motivation "concerned with performance in competition with some standard of excellence" (15, p. 111). Need Affiliation, or motivation to be socially accepted, is conceived as "a disposition to move toward others to elicit positive affective responses from them" (3, p. 409).

At another time, the *S*s had also filled out the SA-S personality inventory[7]. This questionnaire consists of eighty items drawn from Cattell's Sixteen P F Inventory and presumably measures two orthogonal personality dimensions: anxiety and social approach or "permeability." "Individuals scoring at the permeable extreme tend to be sensitive to stimuli, socially oriented, flexible, and imaginative" (22, p. 324).

Although these personality data had been obtained before the experiment, it had been impossible to evaluate them in time to match *S*s in the two conditions. However, subsequent tests showed no significant differences between conditions on any personality measure or intelligence, as measured by scores on the American Council on Education test (ACE).

RESULTS AND DISCUSSION

Differences between conditions. The first step in testing the hypothesis was to ascertain that the experimental manipulations had been effective. A large difference between conditions

[7] These personality data were graciously made available by Neil A. Carrier and John W. Atkinson.

TABLE 1
NUMBER OF CONFORMING RESPONSES PER SUBJECT IN TWO EXPERIMENTAL CONDITIONS

Condition	Number of Conforming Responses													*N*	Mean
	0	1	2	3	4	5	6	7	8	9	10	11	12		
Full Conflict	—	—	9	3	1	5	2	2	—	1	1	2	—	26	4.7
Reduced Conflict	3	6	3	2	1	1	4	2	1	—	—	1	—	24	3.5

on *Ss*' postexperimental ratings of their ability for the task indicated that the simulated practice had indeed succeeded in structuring the situation differently for the two conditions. *Ss* in the Reduced Conflict condition, who had had the initial practice, rated themselves significantly more often as "above average" than the Full Conflict *Ss* ($p < .001$).

Comparing the two conditions on the dependent variable, conformity, we find that the difference is in the predicted direction (see Table 1). However, because of the large variation within conditions, the distributions overlap each other considerably. They are also positively skewed. In order to test the significance of the mean difference, a transformation of the raw data to *arc sin* $\sqrt{\%}$, as given by Snedecor (23), is appropriate, since the scores represent not true measurements but sums of data in zero-one form (independence of conformity per trial). The resulting t of 1.81, with 48 *df*, is significant beyond the .05 level (one-tailed). The null hypothesis of no better performance (i.e., less conformity) in the Reduced Conflict condition can be rejected with reasonable certainty.

Another way of looking at the difference between conditions is given in an analysis by trials. The proportions of conforming responses on each critical trial are shown in Figure 1. On only two trials is the difference in conformity between Full and Reduced Conflict condition negative. A simple sign test shows that the smaller proportion of conforming responses per trial occurred significantly more often in the Reduced than in the Full Conflict condition ($p < .02$, one-tailed). A Wilcoxon signed rank test increases the significance of the difference beyond the .01 level.

In addition, Figure 1 shows that on the *first two* trials only a few *Ss* in either condition conformed. The direction of the (not significant) difference between the two conditions was even reversed. This result, which was not predicted, helps to reject an otherwise serious objection to the procedures employed. It indicates that the overall group difference is not due simply to a better adaptation of the Reduced Conflict *Ss* to the *physical* features of the task as a consequence of their prior practice with a similar one. Apparently, the effects of the majority response reached their real impact on the Full Conflict *Ss* only on the third critical trial, while the majority of the Reduced Conflict *Ss* was still able to withstand them.

The subsequent fluctuations from trial to trial are probably due, in part, to differences between stimuli. These effects are hopelessly confounded in the present design with effects of the spacing of neutral and critical trials[8]. Yet there appears to be a general trend toward decreasing differences between conditions later in the experiment, presumably because the effects of the differential treatment were wearing off. A chi-square test of low vs. high conformity by condition for each half of the experiment results in a significant difference in the predicted direction ($p = .025$, one-tailed) in the first half. In the second half, the difference is no longer significant (although by a parametric test it still is). This decrease of the differential effect is not too surprising if we recall that the treatment actually consisted of a rather artificial distortion of the true facts.

A comparison of conditions on the uncertainty ratings indicated a tendency toward lower mean uncertainty for the Reduced Conflict condition ($.15 > p > .1$) failing to reach significance. This lack of difference had been expected, since the scale points were not well enough defined to ensure interindividual comparability. *Intra*-individual analysis of the different response categories, however, showed

[8] Such effects of the spacing of neutral and critical trials appeared more clearly in the experiment by Asch (2) and one by Olmstead and Blake (18), although in the latter case they were explained as due to differences in the amount of error of the majority judgments.

that most *Ss* in both conditions had rated their conforming responses on the average as more likely to be wrong than their correct responses in agreement with the majority on the neutral trials. This significant difference ($p < .001$ with a sign test) supports the assumption underlying the experiment, that *Ss* would not conform simply because the visual stimuli by themselves were too ambiguous. No consistent differences appeared in a comparison of uncertainties for conforming and independent responses.

The remaining questionnaire responses showed lower ratings for the difficulty of the task in the Reduced Conflict condition, falling short of significance ($.1 > p > .05$), and no differences in the ratings of the *Ss*' visual acuity and the influence of the majority responses. In answer to this last question, about half the *Ss* in either condition insisted that they had not been influenced at all by the majority. Their insistence decreased notably when they became aware, during the discussion, that precisely this effect of the majority responses had been the focus of the *E*'s attention.

Beyond this, the interview did not produce any clear-cut results. Although reactions varied considerably, they all indicated that the *Ss* understandably did not have a clear and consistent picture of the past happenings. Most of them would insist that they had been right whenever they had disagreed with the majority. But in the next sentence they would admit that the judgment of four people had a much higher chance of being correct than their individual one, and therefore they probably had been wrong.

Personality correlates of conforming responses. The large variation within both conditions indicated that an analysis of the personality data might be fruitful. On the basis of earlier results (15), need Achievement was expected to be negatively related to conformity. The meaning of the other personality measures available—need Affiliation, social approach, and anxiety—was not specific enough to warrant any predictions.

While the details of this analysis and its tentative interpretation are reported elsewhere (19, 20), the major results will be briefly described here. A correlational analysis showed that, while there was no relation of conformity to need Affiliation and anxiety or to intelligence (ACE), the other two personality measures *interacted* with the experimental treatment. In the Full Conflict condition, both need Achievement and social approach were *negatively* correlated with conformity (although not significantly). In the Reduced Conflict condition, however, both measures showed a *positive* relation to conformity, which for need Achievement was significant at the .05 level ($r = +.45$, $N = 22$). The *difference* between the two conditions with respect to the correlations with conformity was significant at the .05 level for need Achievement, and the .02 level for social approach. The fact that the two measures were obtained from quite different and experimentally independent sets of data (projective stories and questionnaire responses, given at different times) strongly suggests that this interaction effect is not spurious.

An analysis of variance indicated that the interaction can entirely be attributed to the (expected) reduction in conforming responses of the medium and low need Achievement and social approach scores in the Reduced Conflict condition. Contrary to expectations, the high scorers on either measure showed no absolute mean difference between conditions; but their relative position was changed from less to more conformity by the shift of the other subgroup means.

While the over-all outcome of this analysis thus confirmed the prediction, the conclusion about the effects of the experimental treatment has to be modified. The reduction of conflict, achieved by means of an implicit *argumentum ad hominem*, decreased the amount of conformity exhibited by the majority of the *Ss*. But for the *Ss* scoring high on need Achievement or social approach, this procedure was not effective. Since there is no indication that these *Ss* failed to perceive the difference between themselves and the accomplices which the experimental manipulation had attempted to create, the explanation seems to be that the assumed reduction of conflict did not actually occur for these individuals. As they were highly motivated to achieve, their strong motivation to be correct may have kept the conflict between their own perceptions and the majority judgments alive. Thus, they were forced to vacillate between independence and conformity, and thereby to conform more than the other *Ss* in this condition, who could

make better use of the cues indicating that the majority might be in error.

No entirely satisfactory interpretation of the results obtained with the personality measures has been found[9]. It should be mentioned, however, that the few related experiments on conformity employing contrasting experimental conditions as well as personality measures have shown similar interactions between condition and personality (10, 12, 16).

Further theoretical considerations. It will be recalled that the majority of the *S*s showed the predicted reduction of conforming responses after they had obtained information enabling them to account for the discrepant majority reports. This information had been implicit in the events of the practice, which employed different stimuli and rather different (shorter) exposure times than the test trials. In this situation, one cannot speak simply of reinforcement of specific responses like the individual norm in the autokinetic situation. Neither was *disagreement* with the majority directly reinforced since in the practice, the accomplices' responses eventually agreed with the *S*'s original one. Of course it can be argued that an internal response of *thinking oneself right* was reinforced (13). However, since generalization of this response along unspecified dimensions must be assumed in order to cover different stimuli and different social situations, the value of this concept seems debatable. The interpretation proposed here is that the strengthening of the *S*s' independence involved cognitive processes, especially the *S*s' inference about *relative* abilities of the observers, and its application to the new conflict situation in which the *S*s found themselves. In other words, cognitive relationships among all three elements of the standard social situation—what Newcomb (17) has called the co-orientation system—were involved.

An exclusive emphasis on cognitive processes, of course, does injustice to the complexity of the total situation. The development of the conflict, and the form that its resolution takes, depends on a number of additional factors. Yet it seems that, at least to some extent, two aspects of the situation are separable. To use the language of communication theory, it can be said that the receiver, confronted with different signals in different channels, has to make the decision as to which of the two is "noise" and which the correct message. And in order to make this decision he needs additional information. Which message he in turn *transmits* depends on additional considerations, if he realizes that the factually correct message may not be "socially correct," and if he expects feedback from the social environment. In fact, these effects may easily react on the first-stage decision, and in the extreme case the *S* may give up the attempt to collect independent information and rely entirely on the social channel.

Yet the thesis of this study was that a satisfactory analysis of conformity requires more than the definition of a phenotypic class of responses coinciding with the modal values of the group. Not only, as Asch (2) has argued, must there be a genuine conflict between the different inputs, which is not present in experiments with ambiguous or carefully equated stimuli. The prediction of independent behavior depends on additional information about the source of the conflict. Simply to demand of the *S* (or to train him) to disregard the socially transmitted information without accounting for it (13) may not be "rational" either, since so much of our knowledge about reality is received through social channels, and "social reality" may not be only a delimitable part of our environment but may, in principle, permeate it throughout.

The large interindividual variation in the present experiment and its relation to motivational variables makes it abundantly clear that a cognitive theory alone is not able to explain the events. The readiness to draw the conclusion that the majority is wrong, as well as the willingness to look for information indicating this possibility, may be, in a wider context, important factors in the determination of conformity or independence.

SUMMARY

Fifty *S*s were given the task of identifying nonsense syllables presented tachistoscopically above threshold duration, while being confronted with wrong responses of a majority of secretly instructed *S*s. In simulated practice trials, half the naive *S*s had been given the impression that they had lower recognition

[9] A more extensive treatment of these results can be found in (19, 20).

thresholds than the other observers. This manipulation was intended to reduce the conflict between *Ss'* perceptions and the majority judgments by implying that the majority might be in error. It was predicted that the amount of conformity in this condition would be less than in the standard condition which provided no information enabling the *S* to reconcile the apparent contradiction.

The prediction was confirmed. Absence of the difference between conditions on the first two trials indicated that the main result was not due simply to the adaptation of the Reduced Conflict group to the physical features of the task.

The large variation within conditions was found to be related to two personality measures: need Achievement and social approach. The *Ss* with high scores on these measures performed rather similarly in both conditions, while the predicted difference was obtained through a great reduction of conforming responses by the *Ss* with medium and low scores.

It was concluded that the creation of a frame of reference permitting the interpretation of the majority responses as in error can reduce the amount of conforming behavior. However, some individuals may not be willing or able to make use of such cues.

REFERENCES

1. ALLPORT, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.
2. ASCH, S. E. *Social psychology.* New York: Prentice-Hall, 1952.
3. ATKINSON, J. W., HEYNS, R. W., & VEROFF, J. The effect of experimental arousal of the affiliation motive on thematic apperception. *J. abnorm. soc. Psychol.*, 1954, **49**, 405–410.
4. BRUNER, J. S., & POSTMAN, L. Perception, cognition, and behavior. *J. Pers.*, 1949, **18**, 14–31.
5. COFFIN, T. E. Some conditions of suggestion and suggestibility. *Psychol. Monogr.*, 1941, **53**, No. 4.
6. CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychologist*, 1955, **10**, 191–198.
7. DEUTSCH, M., & GERARD, H. B. A study of normative and informational social influences upon individual judgment. *J. abnorm. soc. Psychol.*, 1955, **51**, 629–636.
8. FRITZ, C. A., JR. Sense-perception and material objects. *Phil. and Phenomenol. Res.*, 1956, **16**, 303–316.
9. GOLDBERG, S. L. Three situational determinants of conformity to social norms. *J. abnorm. soc. Psychol.*, 1954, **49**, 325–329.
10. HARDY, K. R. The influence of affiliation motivation and social support upon conformity and attitude change. Unpublished doctor's dissertation, Univer. of Michigan, 1955.
11. KELLEY, H. H. Two functions of reference groups. In G. E. Swanson, T. M. Newcomb, & E. L. Hartley (Eds.) *Readings in social psychology.* (3rd ed.) New York: McGraw-Hill, 1952.
12. KELMAN, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. abnorm. soc. Psychol.*, 1950, **45**, 267–285.
13. LASKIN, EVA R. Several methods of preparing individuals to resist social pressure to conform. Unpublished doctor's dissertation, Yale Univer., 1955.
14. LUCHINS, A. S. Social influence on perception of complex drawings. *J. soc. Psychol.*, 1945, **21**, 257–273.
15. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.
16. MOUTON, JANE S., BLAKE, R. R., & OLMSTEAD, J. A. The relationship between frequency of yielding and the disclosure of personal identity. *J. Pers.*, 1956, **24**, 339–347.
17. NEWCOMB, T. M. An approach to the study of communicative acts. *Psychol. Rev.*, 1953, **60**, 393–404.
18. OLMSTEAD, J. A., & BLAKE, R. R. The use of simulated groups to produce modifications in judgment. *J. Pers.*, 1955, **23**, 335–345.
19. SAMELSON, F. Group pressure and incongruity in the cognitive field as determinants of conformity. Unpublished doctor's dissertation, Univer. of Michigan, 1955.
20. SAMELSON, F. The relation of achievement and affiliation motives to conforming behavior in two conditions of conflict with a majority. In J. W. Atkinson (Ed.), *The assessment of human motives.* Princeton: Van Nostrand, in press.
21. SHERIF, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, No. 187.
22. SMITH, D. E. P., & WOOD, R. L. Verbal satiation and personality. *J. abnorm. soc. Psychol.*, 1956, **52**, 323–326.
23. SNEDECOR, G. W. *Statistical methods.* Ames: Iowa State Coll. Press, 1946.

Received October 8, 1956.

THE REPRESSION RESPONSE TO IMPLIED FAILURE AS A FUNCTION OF THE HYSTERIA-PSYCHASTHENIA INDEX[1]

CHARLES B. TRUAX

University of Wisconsin

This research was designed in order to test the Freudian theory of repression proper and its relation to scores on the repression index of the Hysteria minus the Psychasthenia scales of the MMPI.

Repression proper (or after-expulsion) is the dissociation or removal from consciousness of anxiety laden material (7). "Removal from consciousness" can be interpreted as simply meaning that the anxiety-laden material is not available for verbalization. Presented in terms of Hullian learning theory (1), the repression hypothesis states that verbal response-producing cues which have been associated with anxiety generate anxiety itself, and that the act of inhibiting the verbal response is learned through the reduction of anxiety. Further, the anxiety should generalize to other verbal responses that are associated with the original anxiety-associated response. Repression, then, should be learned to verbal responses that are associated with the original anxiety-producing response.

If repression is to be regarded as only one of the possible learned responses or mechanisms used to reduce the anxiety drive, however, then not all persons subjected to anxiety may be expected to react by repressing the anxiety-laden material. Erickson (3, 4) has suggested that scores on the Hysteria scale minus scores on the Psychasthenia scale of the MMPI differentiate subjects (*S*s) that characteristically react to anxiety with the response of repression from those who usually respond in some other manner. Empirically, he found that a high *Hy-Pt* group recalled fewer incompleted tasks and were rated as showing more repression than a low *Hy-Pt* group.

Specifically, then, it is hypothesized that those *S*s scoring high on this repression index show more repression responses to anxiety than do *S*s scoring low on it.

In an earlier study by Zeller (15), differential retention for two separate groups following failure-induced anxiety was demonstrated. This finding, however, can be satisfactorily explained in terms of a decremental effect of anxiety upon performance in a complex task as proposed by Farber (6).

The test of repression used in the present research involved the determination of whether, once new learning was established, retention for associations involving anxiety-laden material was poorer than for associations involving neutral material *for the same S at the same time*, and whether, when anxiety was removed, there was no differential rentention of the same materials.

In a recent study of the effects of anxiety and failure, Truax and Martin (13) found an interaction between degree of *personalization* of failure and manifest anxiety. To avoid this problem in research on failure effects, the present study employed an *implied-failure* technique, involving a minimum of personal interaction between experimenter (*E*) and *S*. This procedure has the additional advantage of reducing the importance of *E* in the failure experience, thus probably lessening the release of *S*'s anxiety in the form of hostility toward *E*. In the present study, then, *E* made no comments to indicate *S*'s failure. Instead, the anxiety-producing task, which was irrelevant to the experimental task, was simply made so difficult that *S* could not perform well. Failure was thus *implied* by the performance itself, and since the task seemed ego-involving, anxiety could be legitimately inferred.

METHOD

There were six parts to the present experiment: (*a*) selection of neutral, emotional, and traumatic words to be used as stimuli in a list of 18 paired associates, (*b*) original learning of the paired associates, (*c*) induction of anxiety for the traumatic words, (*d*) first relearning and recall of all paired associates, (*e*) removal of induced anxiety for the traumatic words, and (*f*) second relearning and recall of all paired associates.

The emotional and neutral stimulus words used in the paired associates were selected from a word associa-

[1] The author is indebted to Dr. J. C. Gilchrist for suggestions and criticisms in all phases of the research.

tion list of 79 items. The traumatic stimuli were taken from the titles of subtests of an "intelligence test" used to imply failure. The response words used in the paired associates were three-letter nonsense syllables with an association value of 53 per cent taken from Glaze's list (9). Since the response words were neutral nonsense syllables, it was anticipated that conscious suppression of the response would be minimized and thus permit a clearer test of repression.

Selection of Neutral and Emotional Words for Each S

Words from the word association list were presented on a Gerbrand Tachistoscope for a duration of .5 seconds to each *S* individually. The S was asked to respond immediately with the first word coming to mind. Reaction times to the nearest .05 sec. and all verbal responses to each word were recorded. Depressing a telegraph key simultaneously started both the tachistoscope and a standard electric clock; releasing the key stopped the clock at the beginning of *S*'s verbal response.

In addition, Galvanic Skin Response (GSR) changes were obtained during the reaction time interval and were recorded by means of a General Electric photoelectric continuous recorder. An a.c. GSR apparatus designed by Grant (10) with dry silver electrodes attached to the fingers was used. Since the size of GSR changes varies with *S*'s base level of resistance, the sensitivity of the measure for each *S* was based upon the square root of conductance at his base level to permit intra-individual comparisons. The actual sensitivity of the GSR measure ranged between 30 and 600 ohms.

The neutral words for a given *S* were then defined as the six words to which the smallest GSR deflections and short reaction times were given, while the emotional words for a given *S* were defined as the six words to which he gave the largest GSR deflections and short reaction times. Long reaction time words were not used for two reasons: (*a*) long reaction times is correlated with association or entropy value, and (*b*) failure to respond immediately results in spuriously large GSRs, posaibly because of anxiety over disobeying instructions to respond quickly. As a result of this selection procedure, each *S* had individualized lists of neutral and emotional words, defined by his own responses to them. In addition, the neutral and emotional words for each group of *S*s were equated on the basis of the Thorndike-Lorge word count (13).

Implied Failure and Traumatic Words

In the session with each *S* between original learning and the first relearning and recall, failure was induced and associated with the traumatic words. Each *S* was given a test booklet with the words "Wisconsin Test of Intellectual Abilities: College Form B" displayed on the upper half of the cover sheet. This test consisted of six subtests prominently labeled with the names to be used as the traumatic words in the paired associate task: ***opposites, numbers, completion, reasoning, arithmetic,*** and ***analogies.***[2] Each subtest consisted of ten items of extreme difficulty. At the beginning of each subtest, *E* read the title and gave a brief but glowing account of the abilities measured by the test, stressing their relationship to general intelligence. The name of the subtest (traumatic word) was repeated three times in the instructions for each subtest. The instructions emphasized speed, and a stop watch was used conspicuously. After each *S* completed only half the items, *E* would click the stop watch and say "Stop." No specific comments about the adequacy or inadequacy of the performance were made.

[2] Available on request from the author.

TABLE 1
EXPERIMENTAL SEQUENCE

Day	Session	Intersession Interval	Operation Involved
1	0	—	Selection of emotional and neutral words
2	I	2–3 wks.	Original learning of all 18 paired associates
3	0	24 hrs.	Anxiety induction by implied failure
4	II	24 hrs.	First relearning and anxiety removal
5	III	24 hrs.	Final relearning under neutral conditions

Disabusal of Failure-induced Anxiety

At the conclusion of the first relearning session, each *S* was asked to recall the "intelligence test" that he had taken the day before. The *E* asked *S* how he thought he had done on the test and how he felt about it afterward. After listening to *S*'s remarks, *E* continued:

> [Actually, that wasn't an intelligence test at all. We are trying to study the effect of failure on verbal learning, so we simply made up a fake test. As you probably noticed, the fake test was so difficult that everyone had to do poorly on it. Usually there aren't any time limits on this sort of test, but to make you feel that you failed we stopped everyone when they were halfway thru. By the way, did you notice that part of the response words that you had to learn were the same as the names of the subtests of the fake I.Q. test? (*E* recorded the responses of *S*) Well, I hope the fake test didn't upset you too much. Tomorrow we'll finally get to the GSR conditioning.]

Learning, Recall, and Relearning

Each *S* learned and relearned by the method of adjusted learning (8) to a criterion of two successive correct anticipations for each paired associate. The stimulus and the response words were printed on different sides of 3 x 5 cards with the nonsense syllables paired with the stimulus words so that a given nonsense syllable was paired with each of the three classes of stimuli equally often. With this method of counterbalancing, any difficulty due to the nonsense syllables was controlled.

The 18 cards containing the paired associates were shuffled between each trial to eliminate serial position effects. The stimulus and then the response words were presented to *S* separately for a two-sec. period. A one-sec. interval between pairs and a five-sec. interval between successive list presentations was used. After two

successive correct anticipations for a given pair, that pair was discarded, and learning or relearning continued with the remaining pairs.

Experimental Sequence

The procedural sequence is presented in Table 1. It will be remembered that *S* was led to expect GSR conditioning on the days following the original learning of the paired associates. By this technique, it was hoped that any inter-session rehearsals of the material would be minimized.

Subjects

Individuals from the upper and lower five per cent on the *Hy-Pt* distribution were selected from 650 students in the beginning psychology course at the University of Wisconsin. These *Ss* were then contacted and asked to serve in a "problem solving" experiment. The first 20 *Ss* from each end of the *Hy-Pt* distribution that were contacted volunteered and were used in the present experiment. Thus, a total of 40 *Ss* served in this study, with seven women and thirteen men in the high *Hy-Pt* group and seven women and thirteen men in the low *Hy-Pt* group. Identical procedures were used for both groups.

RESULTS

The performance scores used as the dependent variable consisted of difference scores obtained by subtracting the recall or the learning-relearning scores on the six neutral paired associates from the scores for either the six emotional or the six traumatic paired associates. Four sets of difference scores resulted: (*a*) traumatic recall difference scores, (*b*) emotional recall difference scores, (*c*) traumatic learning-relearning difference scores, and (*d*) emotional learning-relearning difference scores.

Four separate analyses of variance for double classification with repeated measures were performed on the data, corresponding to the four types of difference scores used as the dependent variable (5). In each case, high and low groups based on the upper and lower five per cent of the *Hy-Pt* index were used as the classes, and measures across the three experimental sessions were used as the repeated measures.

The analysis of variance for the recall difference scores and for the emotional learning-relearning difference scores revealed no significant differences. In the case of the analysis for the recall difference scores, the within *Ss* term also proved to be nonsignificant. This lack of individual variation appears to be due to the constricted range of scores caused by relearning. The meaningfulness of these null results with recall scores is, therefore, hard to assess.

TABLE 2
ANALYSIS OF VARIANCE FOR RELEARNING OF TRAUMATIC WORDS

Source	*df*	Mean Square	*F*
High & low groups *Hy-Pt* index	1	130.34	—
Between *Ss* within groups	38	321.07	3.70***
Sessions I, II, III	2	651.33	7.52***
High & low groups × Sessions	2	1,601.74	18.47***
Sessions × *Ss* within groups	76	86.72	
Total	119		

*** Significant beyond the .001 level of confidence.

The results of the analysis of the learning and relearning of the traumatic words in comparison to the neutral words is shown in Table 2. The main effect of learning-relearning sessions is significant beyond the .001 level of confidence, indicating that retention of traumatic words for all *Ss* changed in relation to the neutral words as a function of the operations of anxiety induction by implied failure and the removal of this anxiety by instructions. That is, the traumatic words, which were harder to learn in Session I, became easier to learn by Session III. Further, the interaction of high and low *Hy-Pt* groups by measures across learning-relearning sessions was significant beyond the .001 level of confidence, indicating that the high and low *Hy-Pt* groups responded differently to anxiety induction and anxiety removal.

To evaluate further the interaction of high and low *Hy-Pt* groups by measures across experimental sessions, the Duncan Multiple Range Test was performed on the data (2). Since extreme groups were used in the present study, the .01 level was taken as the acceptable level of significance. The results of this test indicated that the high and low groups did not differ on original learning or on the final relearning of the traumatic words as compared to the neutral words, but did differ in the relearning session after anxiety had been induced (Session II). In Session II, when repression should be manifest, the high group took more trials to relearn the traumatic than the neutral material as compared to either original learning or final relearning. The low

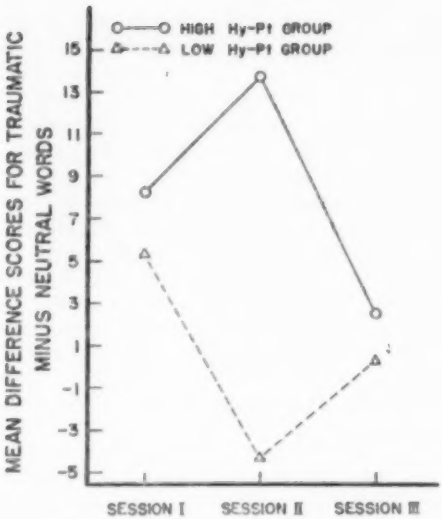

FIG. 1. MEAN DIFFERENCE SCORES FOR TRAUMATIC WORDS

group, on the other hand, took fewer trials to relearn the traumatic than the neutral material in Session II as compared to original learning. The difference between Session II and Session III for the low group, however, fell short of the required significance level ($p = .06$).

In Fig. 1, the mean difference scores for the high and low *Hy-Pt* groups for the traumatic words are plotted as a function of the sessions. Each point represents the trials to the criterion taken by 20 *Ss* on the six traumatic minus the six neutral words. It will be remembered that Session I was a neutral condition of original learning, Session II the relearning after anxiety induction by the technique of implied failure, and Session III the final relearning after the removal of the anxiety by instructions.

DISCUSSION

The Learning-Relearning Measure of Repression

When the operation of implied failure is taken as the source of anxiety, the results of the present study are consistent with the Freudian theory of repression proper and the hypothesis that the *Hy-Pt* index differentiates repressors from non-repressors. As shown in Fig. 1, the high group tended to show repressive responses to anxiety, while the low group tended to be facilitated in their performance by anxiety. That these findings were not the result of supression, rather than repression, is indicated by the answers to the question: "Did you notice that part of the response words that you had to learn were the same as the names of the subtests of the fake IQ test?" All *Ss* were quite surprised when this relationship was pointed out to them. If one grants that there was no reason for the *S* to lie at this point—or even that *Ss* are not good actors—then the existence of the class of traumatic words was not known by the *Ss*. It is felt that the 24-hr. interval before and after anxiety induction was crucial to the present study. If the failure had occured in temporal juxtaposition to the paired-associate task, it seems likely that many *Ss* would have "seen through" the experiment.

Unfortunately, the results, although showing significant differences in repression, are not free from confounding of the method used with the act of repression. That is, the amount of retroactive inhibition produced by the method of adjusted learning is not known. Since the traumatic words tended to be more difficult in original learning, they were not only presented on more trials, but they were also learned *last*. This, then, allowed for more retroactive inhibition of the neutral words which would tend to make for better relative retention of the traumatic words in Session II. Thus, since repression was not operative in the low group to counteract the effect of retroactive inhibition, these *Ss* showed faster relearning of the traumatic words in Session II as compared to the neutral words. In Session III, of course, retroactive inhibition would act on the traumatic words since they were relearned first in Session II. If this situation actually did occur, there is no need to postulate an antithetical process such as vigilance to account for the performance of the low group. Whether the results obtained with the traumatic words are due to a process of vigilance in the low group and of repression in the high group or, instead, to a combination of repression and the effects of retroactive inhibition inherent in the method of adjusted learning can only be tested by further research.

When emotional words are taken as the source of anxiety, the lack of significant differences between high and low groups in the retention of anxiety-laden material is

not consistent with past research where differences have been found in relearning scores (14). Such positive results may have been less a function of the effects of anxiety associated with emotional words than of the use of long reaction times as an indicator of the words' emotionality. Emotionality defined in terms of long reaction times tends to be related to association value, which is also measured by reaction time, or to entropy value. Thus, long reaction times indicate low association values and short reaction times high association value for the individual *S*. Studies reporting differences in retention of emotional and neutral words and using reaction time as a measure may have confounded the effects of anxiety and association value. Since no difference in the retention of emotional words was found in the present study where emotional and neutral words were equated in association value by means of (*a*) reaction time for each *S* and (*b*) the Thorndike-Lorge word count, this methodological point merits further attention.

The Recall Measure of Repression

The design of the present experiment did not seem adequate to allow for a clear test of retention of anxiety-associated words in terms of recall scores. A range of only seven scores was possible, and the effect of relearning reduced the observed individual differences to nonsignificance. This aspect of the hypothesis, then, must be regarded as essentially untested.

Implications of the Effect of Implied Failure

The results indicate that anxiety does develop in a situation where *S* perceives himself as not doing well, even though no knowledge of the adequacy of his performance is given to him by *E*. This suggests that, in studies of learning designed to evaluate differences in the difficulty or complexity of a task, there exists the very real danger of confounding the effects of task complexity with the effects of failure-induced anxiety. It is common for *S*s to discuss an experiment with others who have already served in it before they volunteer to participate. If the experienced *S* has served in a simpler task and the naive *S* is to serve in a more complex one, then the experimental situation may imply failure to the naive *S* on the basis of his false performance norms. The resulting confounding of task complexity with the effects of failure-induced anxiety should be particularly prevalent in such studies as those of concept-formation, where the nature of the task remains unchanged but the problem increases in complexity. Thus, for example, the presence of irrelevant information might result in greater time consumption while the occurrence of failure-induced anxiety might result in an inability to complete the task successfully.

To control for the differential presence of implied failure in different levels of task complexity, it is suggested that (*a*) generous performance norms be given to each *S*, (*b*) each *S* be told that different *S*s are given tasks differing in difficulty, and (*c*) questionnaires be given to *S* at the conclusion of the experiment to evaluate the differential presence of implied failure in groups where task complexity differs.

The Relation Between the Present Results and Manifest Anxiety

Previous findings, like those reported by Zeller (15), have been susceptible to interpretation in terms of simply the deteriorating effects of failure-induced anxiety. No repression hypothesis seems necessary. In the present study, however, where each *S* acted as his own control and the *Hy-Pt* groups differed significantly in their response to anxiety, a repression process appears to be clearly demonstrated.

The present results cannot be handled by a drive-anxiety interpretation alone. Since scores on the Manifest Anxiety Scale were available for the *S*s used in the present research, the *S*s could be divided into high and low anxiety groups. The high *Hy-Pt* group is a low anxiety group (all scores less than 13), while the low *Hy-Pt* group is a high anxiety group (all scores greater than 23). Previous results on the interaction of individual differences in MAS scores and failure anxiety are consistent with the notion that anxiety acts as a drive (6, 13). Considering only the drive aspect of anxiety in the present study, however, would lead to the prediction that the high anxiety group would perform more poorly than the

low anxiety group on Session II (where anxiety is present). The results are in exactly the opposite direction.

While the present results are not at all a contradiction of the drive-anxiety hypothesis, the results do demand that more attention be placed on habits associated with anxiety. It is suggested that the Hysteria variable[2] is a measure of the specific habit of repression as a response to anxiety. Viewed from this perspective, the results are consistent with the clinical notion of repression and with current Hullian learning interpretations of the interaction of drive and habit.

SUMMARY

The present research was carried out in order to study the effects of anxiety induced by *implied-failure* in relation to the repression hypothesis and the personality factor involved in the Hysteria scale minus the Psychasthenia scale of the MMPI.

A total of 40 *Ss* served in the experiment: 20 *Ss* from the upper five per cent of the *Hy-Pt* distribution and 20 *Ss* from the lower five per cent. There were six parts to the experimental procedure: (*a*) selection of six neutral, six emotional, and six traumatic words to act as stimulus words, and 18 medium association value nonsense syllables to act as response words for paired-associate learning, (*b*) original learning of the 18 paired associates to a criterion of two successive correct anticipations per pair, (*c*) induction and association of anxiety with the six traumatic words, (*d*) recall and relearning of the 18 paired associates, (*e*) removal of the anxiety associated with the traumatic words, and (*f*) final recall and relearning of the 18 paired associates after removal of anxiety associated with the traumatic words.

The dependent variables used consisted of discrepancy scores obtained by subtracting the trials to reach criterion for the neutral words from trials to criterion for either the emotional or traumatic words. Thus, each *S* served as his own control.

The results were interpreted as consistent with the repression hypothesis and with the hypothesis that the *Hy-Pt* index is a measure of repressors. The results were also discussed in relation to the drive-anxiety hypothesis and in relation to the implied-failure technique.

[2] Little and Fisher (11), in a cluster analysis of the Hysteria Scale found two orthogonal factors; denial of symptoms and admission of symptoms. In rescoring the subscales for the high and low groups used in the present study, it was found that the admission subscale contributed 53% of the Hysteria score for the nonrepressors but only 6% of the Hysteria score for the repressors.

REFERENCES

1. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
2. DUNCAN, J. Multiple range and multiple F tests. *Biometrics*, 1955, **2**, 1–44.
3. ERIKSEN, C. W. Defense against ego-threat in memory and perception. *J. abnorm. soc. Psychol.*, 1952, **47**, 230–235.
4. ERIKSEN, C. W. The case for perceptual defense. *Psychol. Rev.*, 1954, **61**, 175–182.
5. EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart & Co., 1950.
6. FARBER, I. E., RUSSELL, W., & ANDREAS, B. The effects of failure upon performance in verbal and motor learning situations: A review of work done at the State University of Iowa. Reported to Midwest. Psychol. Ass., April, 1949.
7. FREUD, S. Repression. *Collected Papers*, Vol. IV. London: Hogarth, 1925. Pp. 84–97.
8. GILLETTE, A. L. Learning and retention; a comparison of three experimental procedures. *Arch. Psychol.* 1936, **28**, No. 198.
9. GLAZE, A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–267.
10. GRANT, D. A. A convenient alternating current circuit for measuring GSR's. *Amer. J. Psychol.*, 1946, **59**, 149–151.
11. LITTLE, K. B., & FISHER, J. The internal structure of the MMPI Hysteria Scale. Paper read at Bay Area VA Clin. Psychol. Res. Meeting, Oakland, Calif., Oct., 1952.
12. THORNDIKE, E. L., & LORGE, I. *A semantic count of English words*. New York: Teachers College, Columbia Univer., 1938.
13. TRUAX, C. B., & MARTIN, B. The immediate and delayed effect of failure as a function of task complexity and personalization of failure instructions. *J. abnorm. soc. Psychol.*, 1957, **55**, 16–20.
14. WORCHEL, P. Anxiety and repression. *J. abnorm. soc. Psychol.*, 1955, **50**, 201–205.
15. ZELLER, A. F. An experimental analogue of repression: II. The effect of individual failure and success on memory measured by relearning. *J. exp. Psychol.*, 1950, **40**, 411–422.

Received October 9, 1956.

INTRA-INDIVIDUAL CONSISTENCY IN CONCEPTUALIZATION

LOUIS J. MORAN
University of Texas

LAURENCE S. McGAUGHRAN[1] AND DON LEVENTHAL[2]
University of Houston

THIS study is concerned with the general problem of whether individuals tend to exhibit a characteristic approach that is consistent and predictable on a variety of conceptual tasks. Both normal and psychiatric groups are used, not for purposes of intergroup comparison, but to examine intra-individual consistency in the conceptual behavior of subjects, whether normal, schizophrenic, or brain-damaged.

For the present purpose, a concept may be defined as a principle under which a variety of particulars may be grouped. These principles may be characterized in a number of ways. For example, a concept may vary in its degree of "publicness," i.e., the degree to which the principle covering a grouping is shared and freely communicated by the majority of persons within the same subculture. Thus, a subject who collects several three-sided figures under the concept *triangles* is using a more "public" principle than a subject who collects the same figures under the relatively private concept, *pleasing*. In other words, a concept may vary in the extent to which it permits public prediction of the objects or events which it is intended to categorize.

Concepts may be classified also in terms of the range or order of objects or events which may be grouped under them. Thus, a subject who characterizes a group of cattle as *organisms* is using a higher order of classification than a subject who uses the concept *animals*; the latter, in turn, is using a higher order of classification than a subject who specifies *cows*. In this instance, the concept *organism* may be called highly "open" in relation to the concept *cows* in terms of the range of other objects that potentially might be included under the principle used.

Concepts may, of course, be classified in a number of other ways. The present analysis, however, involves only the two variables mentioned, "publicness-privateness" and "openness-closedness." This schema delineates four "conceptual areas" which are represented by the quadrants formed through the intersection of the two variables. Thus, concepts may be classified as public-open, public-closed, private-open, or private-closed. It has been shown that by means of this schema, such groups as the brain-damaged, the schizophrenic, and the intellectually bright and dull may be reliably differentiated (1, 2, 3).

The present study is designed to test the hypothesis that individuals evidence a consistent tendency to operate within a specific quadrant in responding to a variety of conceptual tasks. Such a consistent tendency constitutes a conceptual trait, implying a general and predictable behavior potential toward the solution of conceptual problems.

METHOD

Subjects

The initial sample consisted of 77 subjects (*Ss*). Of these, 40 were patients in the Houston Veterans Administration hospital, undergoing treatment for or recuperating from nonpsychiatric illnesses which would not affect in any known way their performance on conceptual tasks. A second group of 37 *Ss* were backward psychiatric patients in the Waco Veterans Administration Hospital, all diagnosed as chronic paranoid schizophrenics.

Procedure

In a study conducted several years ago by the senior author (4), a battery of seven word-meaning tests was administered to these *Ss* in order to test several hypotheses concerning the schizophrenic's understanding and use of word meanings. Their performance on two of these tests, a 25-item synonym test and a 17-item similarities test, is reanalyzed here. The new scoring procedures, based upon the conceptual-area schema, were derived a priori. Each item was scored on every test before proceeding to the next item. The consistency with which each *S* used a given conceptual area on these two tests and on an object sorting test, which was administered at the same time, was tested.

Sorting test. The sorting task was administered individually in accordance with the directions given by

[1] Employed part-time at the Houston V. A. Hospital where part of this study was conducted.

[2] Now on the staff of the Topeka State Hospital.

Rapaport (5). In brief, the technique involved the sorting of 33 objects familiar in everyday experience. One part consisted of seven tasks requiring *S* to group with some particular object the others that "belong with it." The second part consisted of twelve tasks requiring *S* to identify a basis for the grouping of a number of objects arranged by the experimenter.

Each of the responses to the sorting test items was assigned to one of the four conceptual areas: public-open, public-closed, private-open, private-closed. Publicness-privateness was scored in terms of the extent to which an individual concept permitted potential public prediction of limits. For example, the boundaries for a concept of "redness" are rather clearly denotable; the conceptual boundaries for "attractiveness" would be considerably less so. The variable of "openness-closedness" was scored in terms of the number of common attributes which were used or implied in each conceptual definition. For example, the classification of a group of objects as "all are red" is "open" relative to such a classification as "all are *round* and *red* and *rubber*." This latter concept binds together more of the perceptible attributes of the objects and, hence, limits the number of other objects that might potentially be included in the same domain. An extended treatment of this scoring system is provided by McGaughran and Moran (2). Interjudge agreement, based upon 55 randomly selected items, was 93 per cent.

Similarities test. The similarities test was composed of 17 groups of two to four words which may be categorized under a single concept; for example, *boat, train, car*: "all modes of transportation." The instructions were as follows: "Now I am going to read you some groups of words. I would like you to tell me what the words in the group have in common; that is, in what way they are alike."

A response was scored as "public" if *S* grouped the words into an appropriate category such as *door-window*, "openings"; or described a meaningful similarity among the words like "both part of a house." A response was scored as "open" if *S* responded with a concept that was of a higher order of classification than any of the individual stimulus words; for example, *door-window*, "openings." Interjudge agreement, based upon 200 randomly selected items, was 86 per cent.

Synonyms test. The synonyms task was given in two parts. The first part consisted of 25 very familar words which were read one at a time to the *S* under the following instructions: "Tell me all the words you can think of that have the same meaning as this word. For example, words having the same meaning as *little* are: small, tiny, diminutive, and so on. Now, how many words can you think of that have the same meaning as the word *house*, etc.?"

Performance on this part of the test placed the *S* somewhere on the public-private variable. Responses to each of the 25 words were scored as follows:

(*a*) Score 2—gave more correct synonyms than incorrect synonyms.
(*b*) Score 1—gave equal number of correct synonyms and incorrect synonyms.
(*c*) Score 0—gave more incorrect synonyms than correct synonyms.

Categories *a*, *b*, and *c* represent levels of communicability in *S*'s responses. The scoring is relatively independent of both *S*'s brightness and the number of responses given. Thus, an *S* may give one correct synonym for each stimulus and still achieve the highest possible score on the public-private variable.

The second part of the synonym task places the *S* somewhere on the open-closed variable. The same 25 words were presented. After each word, however, eight choice words were listed. The number of correct synonyms varied from two to five of the eight alternates. The incorrect alternates were associated in some way with the test word but were not synonyms for it, e.g., house-roof. Subjects were given the following instructions: "After each capitalized word in this list are a number of words in small print. Underline each of the words in small print that has the same meaning as the word in capital letters. Do not guess."

In effect, for each word *S* was given eight concepts to "sort" and to place those together that belonged with the stimulus concept on the basis of synonymity. If he grouped most or all of the correct synonyms, he utilized to the maximum the concept of synonymity and was relatively "open." If he grouped a few of the correct synonyms, he was restricting the concept of synonymity and was relatively closed in his sorting. To the extent that nonsynonyms were also grouped, his concept of synonymity was overinclusive and, hence, was relatively open. A measure of open-closed, then, was given by simply totalling the number of concepts underlined.

In short, each *S* performed on three different conceptual problems. The first task involved the formation by *S* of conceptual groupings with concrete objects. The second task required the formation of a concept to serve as the principle for grouping words. The third task served as a more intensive appraisal of *S*'s use of a single principle, that of synonymity. The degree of consistency in approach to these three tasks was estimated by Kendall's coefficient of concordance (W_c).

RESULTS

Individuals in both groups were reasonably consistent in the nature of their solution on the three different conceptual tasks. On the public-private variable the normal group achieved a W_c of .47 ($p < .05$) and the schizophrenic group a W_c of .67 ($p < .01$). On the open-closed variable the W_c's were .57 ($p < .01$) for the normal group and .63 ($p < .01$) for the schizophrenic group.

The public-private and open-closed variables appear to vary independently. Rank-order correlations between these two variables did not approach significance with either the normal group or the schizophrenic group.

FOLLOW-UP STUDY

In an independent study of differences in conceptualization between schizophrenic *S*s and brain damaged *S*s, the same measures were administered to 30 chronic paranoid

schizophrenic patients and to 30 patients with a diagnosis of generalized brain damage at the Topeka State Hospital. At the time the protocols were scored, the scorer was not aware of the hypothesis of intra-individual consistency. Coefficients of concordance were determined after analyses of intergroup differences were completed.

On the public-private variable, the schizophrenic group achieved a W_c of .78 ($p < .01$) and the brain damaged group a W_c of .67 ($p < .01$). On the open-closed variable, the W_c's for the schizophrenic group and the brain damaged group were .48 ($p < .05$) and .63 ($p < .01$), respectively. In both groups, the public-private and the open-closed variables were uncorrelated with each other.

DISCUSSION

In previous studies using the conceptual-area schema, inter-individual comparisons have shown wide differences in approach to conceptual problems. The measurement of individual differences has proved serviceable particularly in the discrimination of deviant groups, such as the psychotic (2) and the brain damaged (3). It was the intent of the present study to test for *intra*-individual consistency on a variety of conceptual tasks. The results of this investigation seem to warrant the conclusion that intra-individual consistency in conceptual-area performance is sufficient to merit the term "conceptual trait."

SUMMARY

An object sorting task, a similarities task, and a synonyms task were administered individually to 40 nonpsychiatric medical patients and to 37 chronic paranoid schizophrenic patients. Performance on each task was scored in terms of a "conceptual area" schema, which placed the subject's conceptual solutions in one of four "areas": public-open, public-closed, private-open, private-closed. Results were evaluated by Kendall's coefficient of concordance. Both nonpsychiatric and psychiatric subjects were found to be consistent in their approach to this variety of conceptual tasks. These results were extended to new samples of paranoid schizophrenic and of brain damaged subjects. It was concluded that the conceptual-area schema provides a means of identifying and measuring conceptual traits.

REFERENCES

1. MCGAUGHRAN, L. S. Predicting language behavior from object sorting. *J. abnorm. soc. Psychol.*, 1954, **49**, 183–195.
2. MCGAUGHRAN, L. S., & MORAN, L. J. "Conceptual level" vs. "conceptual area" analysis of object-sorting behavior of schizophrenic and nonpsychiatric groups. *J. abnorm. soc. Psychol.*, 1956, **52**, 43–50.
3. MCGAUGHRAN, L. S., & MORAN, L. J. Differences between schizophrenic and brain-damaged groups in conceptual aspects of object sorting. *J. abnorm. soc. Psychol.*, 1957, **54**, 44–49.
4. MORAN, L. J. Vocabulary knowledge and usage among normal and schizophrenic subjects. *Psychol. Monogr.*, 1953, **67**, No. 20 (Whole No. 370).
5. RAPAPORT, D. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945.

Received October 8, 1956.

DISTINGUISHING CHARACTERISTICS OF COLLABORATORS AND RESISTERS AMONG AMERICAN PRISONERS OF WAR

EDGAR H. SCHEIN, WINFRED F. HILL, HAROLD L. WILLIAMS, AND ARDIE LUBIN[1]

Walter Reed Army Institute of Research

AMERICAN soldiers captured during the Korean conflict collaborated with the Communists in varying degrees. Most prisoners of war (POWs) committed trivial acts such as signing peace petitions, but a smaller number of men engaged in persistent collaboration which included writing, signing, and soliciting signatures for peace petitions, delivering anti-United Nations lectures to other POWs, aiding in the preparation and distribution of communist propaganda, encouraging fellow POWs to collaborate, informing on other POWs, and generally aiding the Chinese in their indoctrination program (10). It was not clear from the accounts of the repatriates themselves whether one should seek the causes of such collaboration in the pattern of stresses which the men faced, or in personality and background factors in the men themselves, or in some combination of these circumstances. Many hypotheses have been proposed, varying in the degree to which they attribute collaboration to the so-called communist "brainwashing" techniques, or blame it on "weaknesses" in the men themselves, or on weaknesses in American society (3, 5, 6, 7, 8, 10).

The present study is an attempt to determine whether factors in the background and personalities of the repatriates are related to collaboration. However, it should be pointed out that the search for determinants in the background and personalities of the men does not represent a lack of emphasis in our thinking on the role of the stressful environment (as described in references 4, 8, 11). Rather, it is felt that whatever personal determinants emerge will illuminate the complex relationship between specific situational stresses and the individual upon whom they impinge.

METHOD

Variables and Instruments

It was decided to investigate background factors, intellectual variables, and personality variables. However, we could not expect to test any man for more than two hours, which resulted in a somewhat arbitrary set of decisions concerning tests and prevented an extensive sampling of personality dimensions.

In order to determine whether such factors as *home background* and *social status* related to collaboration, the following biographical information was obtained from each man: (*a*) military rank, (*b*) whether the subject was in the regular army, reserves, National Guard, or had been drafted, (*c*) location of pre-service home community, (*d*) age, (*e*) years of education, (*f*) civilian occupation, (*g*) length of army service prior to capture, (*h*) length of internment, (*i*) religion, (*j*) whether the subject had spent most of his life in a large city, small city, small town, or farm, and (*k*) whether he had spent most of his life with both parents, one parent, or neither parent.

In order to determine whether level of *intelligence* related to collaboration, the information and comprehension subtests from the Wechsler-Bellevue Adult Intelligence Scale and the two halves (vocabulary and abstraction) of the Shipley-Hartford test (years of education may also be included with these intellectual variables) were administered.

The selection of suitable *personality* dimensions presented a more difficult problem. Because collaboration is probably a very complex multidetermined behavior pattern, specific hypotheses were not formulated. However, it seemed likely that a man's relationship to authority figures should play a large role in determining his response in the prison situation. On the one hand, it might be expected that opportunistic disregard for *any* form of authority would produce a pattern of collaboration. On the other hand, a rigid dependency on *any* form of authority might also lead to collaboration if the POW accepted his Chinese captor as the primary source of authority. These two possibilities led us to administer the Psychopathic Deviate Scale of the Minnesota Multiphasic Personality Inventory (MMPI) and an authoritarian scale resembling the California F scale (1). The F scale could not be used because of the lower educational level of the subjects, and because a number of the F scale items involve content which we considered to be too emotion laden for the POWs (our items put primary emphasis on intolerance of ambiguity). Our scale and the F scale were subsequently administered to a sample of 74 army enlisted men (not former POWs) and were found to correlate +.81 with each other.[2]

[1] At the time of the study, Capt. Schein, Capt. Williams, and Sp. 3 Hill were at the Walter Reed Army Institute of Research. Dr. Schein is now at the Massachusetts Institute of Technology and Dr. Hill is at Harvard University.

[2] To save printing costs, this authoritarian scale has been deposited with the American Documentation Institute. Order Document No. 5298 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting

TABLE 1
PER CENT OF REPATRIATES IN THE THREE CRITERION GROUPS

Group	Test Sample		Total Population	
	N	Percent	*N*	Percent
Resisters	39	5.1	138	4.9
Neutrals	608	80.1	2235	78.6
Collaborators	112	14.8	470	16.5
Total	759	100.0	2843	100.0

As an additional general measure the short form of the MMPI was administered, based on a factor analysis of the full MMPI by Welsh (2, 12). This scale permitted us to obtain a general index of anxiety or neuroticism (Welsh's *A* scale) and a general index of repression, social inhibition, and denial (Welsh's *R* scale). The biographical questions, intelligence test items, and personality scale items were mimeographed for group administration.

Subjects

The instruments just described were administered to a sample of 759 repatriated army POWs (22 officers and 737 enlisted men), or about one-fourth of the total 2843 army POWs repatriated. These men were tested in groups of 10 to 15 on board three ships bound from Korea to the United States. As far as possible, techniques of systematic randomization were used in selecting the ships and selecting the sample of men on each ship. Subjects (*Ss*) were told that the testers were interested only in general information about the ordeal of imprisonment and the kind of people who were able to survive it, not in the responses of particular individuals. Testing conditions were not ideal because of crowded quarters, occasional seasickness, interruptions, and a general state of guardedness in the *Ss*. However, in the main, motivation was considered adequate by the psychologists administering the tests on each of the ships.

Criterion of Collaboration

The criterion information regarding collaboration was obtained from the army. Only those men were listed as collaborators against whom the army had reliable deleterious information and whose cases warranted some form of disciplinary action. It is quite possible, therefore, that some of the men classified here as noncollaborators or "neutrals" actually had collaborated to some extent. It should be noted that this classification is based on overt collaborative behavior and does not necessarily imply any ideological conversion.

A few of the noncollaborators were classified as "resisters" on the basis of their prison camp behavior. This group included those men whose army dossiers indicated that they were considered by their Chinese captors as "reactionaries" and were segregated from other POWs, and those men who were recommended by the army for decorations for their behavior in prison camp.

Thus, the final criterion consisted of three categories: (*a*) the resisters, (*b*) the neutrals (those who neither resisted nor collaborated), and (*c*) the collaborators. The distribution of these three criterion groups is shown in Table 1. As can be seen, this distribution closely approximates the distribution of these three groups in the total population of released army POWs. No significant biases appear to have been introduced by the sampling procedures.

Analysis

An analysis of variance was performed on each of the quantitative variables. From each of these there was obtained both an eta coefficient, indicating the degree of relationship between the given variable and the criterion, and an *F* ratio, from which the significance level was obtained. Whenever this was significant, *t* tests were performed to determine whether any of the criterion groups differed from each other. Because of the unequal numbers of *Ss* in these groups the *t* test procedure was used, in spite of the fact that the probability levels can not be interpreted exactly. However, the fact that we have only three groups to compare reduces the likelihood of a marked bias by the use of this technique.

Chi squares were computed for all the variables which were in categorical form (nonquantitative).

RESULTS

The results on all the quantitative variables are shown in Table 2. The most striking feature of these results was that the resisters and collaborators always deviated from the neutrals in the same direction. Moreover, in every case in which the relationship was significant, both of these extreme groups were higher on the variable than the neutrals.

It will be observed that the biographical variables are the most discriminating: all three are significant at the .001 level. Intellectual variables are fairly discriminating, and all but education are significant. Only one personality variable is significant, the *Pd* scale of the MMPI, but its relationship to the criterion is higher than that of any other variable except length of internment. Scores on the authoritarian scale fail completely to discriminate the groups.

Although many of these variables are highly significant (frequently far beyond the cutoff for the .001 level), the eta coefficients are uniformly small. The most powerful predictor actually accounts for less than four per cent of the variance of the criterion. Hence, for *practical* purposes this is essentially a catalogue

$1.25 for 35 mm. microfilm or $1.25 for 6 by 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

of negative results: a number of variables discriminate significantly, but none of them discriminates importantly.

Moreover, between any two criterion groups, even the statistical significance is not striking. Of the eight variables that show overall significance, only five significantly differentiate resisters from neutrals, and six differentiate collaborators from neutrals. The significance levels of these comparisons are shown in the last two columns of Table 2. No comparisons between resisters and neutrals are shown, as none of these was significant.

Among the *categorical* variables, all of which were of the biographical type, only one reached even the .05 level of significance when tested with chi square. This was type of enlistment, in which the same pattern appeared as in the quantitative variables, there being a higher percentage of both resisters and collaborators among the regular army men than among the draftees (officers, reservists and, National Guardsmen were not included in this analysis). Neither army rank, civilian occupation, religion, number of parents present in the home, rural vs. urban background, nor location of home discriminated the criterion significantly. Even the one significant qualitative variable, type of enlistment, was probably artifactual, as will appear later.

TABLE 2

RELATION OF QUANTITATIVE VARIABLES TO CRITERION

Variable		Criterion Group			Eta	Significance Level		
		R*	N*	C*		F ratio	R vs. N	C vs. N
Length of internment (months)	$\bar{X}$	31.0	26.6	31.9	.19	.001	.001	.001
	s	5.9	11.0	5.9				
Length of service (months)	$\bar{X}$	50.4	30.3	35.5	.15	.001	.05	NS
	s	57.8	28.1	37.0				
Age (years)	$\bar{X}$	27.0	24.1	25.2	.16	.001	.05	.05
	s	9.1	5.9	6.8				
Education (years)	$\bar{X}$	9.5	9.2	9.5	.06	NS	NS	NS
	s	2.5	2.0	2.1				
Information score	$\bar{X}$	14.4	12.3	13.3	.15	.001	.001	.02
	s	3.5	3.7	3.9				
Comprehension score	$\bar{X}$	10.4	9.2	9.7	.09	.05	NS	NS
	s	3.9	3.2	2.8				
Vocabulary score	$\bar{X}$	22.8	20.5	23.1	.13	.001	NS	.001
	s	8.0	7.3	7.6				
Abstractions score	$\bar{X}$	20.5	19.8	23.1	.11	.01	NS	.01
	s	10.3	9.7	10.2				
MMPI *Pd* scale	$\bar{X}$	18.2	16.7	18.9	.19	.001	.05	.001
	s	4.3	3.9	5.4				
MMPI *A* scale	$\bar{X}$	11.2	9.8	11.0	.08	NS	NS	NS
	s	7.4	6.5	6.7				
MMPI *R* scale	$\bar{X}$	16.3	16.9	16.8	.03	NS	NS	NS
	s	5.1	5.7	4.7				
Authoritarian scale	$\bar{X}$	25.1	25.5	25.1	.03	NS	NS	NS
	s	4.1	5.1	4.9				

* R stands for resisters, N for neutrals, and C for collaborators.

TABLE 3

RELATION OF QUANTITATIVE VARIABLES TO CRITERION AMONG MEN INTERNED 31 MONTHS AND OVER

Variable		Criterion Group			Eta	Significance Level		
		R	N	C		F ratio	R vs. N	C vs. N
Length of service (months)	$\bar{X}$	57.3	33.4	41.7	.19	.001	.05	NS
	s	61.5	26.7	41.1				
Age (years)	$\bar{X}$	28.0	24.4	25.6	.22	.001	.001	.05
	s	5.2	3.9	5.0				
Education (years)	$\bar{X}$	9.4	8.8	9.5	.13	.01	NS	.01
	s	2.4	1.8	2.2				
Information score	$\bar{X}$	14.6	12.0	13.5	.21	.001	.001	.01
	s	3.6	3.5	4.0				
Comprehension score	$\bar{X}$	10.6	9.0	9.8	.15	.01	.05	.05
	s	4.2	3.0	2.9				
Vocabulary score	$\bar{X}$	23.3	19.6	22.9	.20	.001	.02	.001
	s	7.2	7.2	8.2				
Abstractions score	$\bar{X}$	19.1	18.6	22.4	.15	.01	NS	.01
	s	10.7	9.3	10.4				
MMPI *Pd* scale	$\bar{X}$	18.8	17.0	18.9	.18	.001	.05	.01
	s	4.6	4.0	5.7				
MMPI *A* scale	$\bar{X}$	11.1	10.2	11.1	.06	NS	NS	NS
	s	7.9	6.5	7.1				
MMPI *R* scale	$\bar{X}$	16.4	17.0	16.6	.04	NS	NS	NS
	s	5.1	4.6	4.9				

Because *length of internment*, which is essentially a situational variable, was so high on the list of predictors, it seemed that a purer measure of the effect of the other variables might be obtained by holding it constant. We therefore re-analyzed a number of variables within the subsample of 457 men interned 31 months and over (i.e., those captured in the first half-year of the war). Table 3 shows the results of this re-analysis. For no variable is the trend changed, and for the majority of variables the magnitude of the relationship is somewhat increased. Education, in particular, now becomes statistically significant. Moreover, for the first time, there appears a significant difference between resisters and collaborators (not shown in the table) on one variable, age. The resisters are older, on the average, than the collaborators. All the etas, however, are still quite small and of no practical value. Scores on the authoritarianism scale were not re-analyzed because of its failure to discriminate in the total sample.

Re-analysis of the one significant categorical variable (enlistment type) in the restricted sample proved to be trivial, because only one per cent of the long-interned enlisted men were draftees, making the use of chi square both impossible and superfluous. The apparent

difference between enlistment types on the criterion is thus easily explained by the longer average internment of regular army men than of draftees.

A number of intercorrelations were found among the predictors. The intellectual variables tended to be positively related both to age and length of service (themselves correlated, of course), and to the *Pd* scale of the MMPI. However, a covariance analysis indicated that the *Pd* scale still discriminates almost as well when any of the intellectual variables is held constant. The information score also remains significant when age is held constant. In view of the small predictive power of all the variables, a more complete analysis of their interrelations did not seem worthwhile.

DISCUSSION

Perhaps the most important finding in this investigation is negative—none of the variables tested predicted the criterion to any useful extent. Thus, our results argue against any simple explanation of resistance or collaboration in terms of easily measured background or personality variables. For *practical* purposes this study might best be summarized by saying that resisters, neutrals, and collaborators do not differ appreciably on any of the variables investigated.

Particularly striking is the failure of our authoritarianism scale to discriminate the groups. Apparently, the kinds of personality dimensions measured by this instrument are not related to prison camp adjustment, unless one assumes that the repatriates were more guarded and inhibited in answering these items than MMPI items. As we stated previously, we deliberately attempted to remove items that might arouse defensiveness, but we have no way of knowing the extent to which this effort was successful.

Similarly, occupation, number of parents, location of home, rural-urban origin, and religion failed to discriminate the groups, though it is possible that the failure to obtain significant differences may be due to the lesser power of the chi-square test.

Nevertheless, a number of statistically significant differences were found, and theoretically these pose interesting problems. As may be recalled, it had been our hunch that collaboration might be related to opportunistic disregard of authority. We hoped to measure this factor by means of the MMPI *Pd* scale and obtain the expected relationship. But why should resisters also have come out high on this scale? This finding suggests that we may have to interpret the *Pd* scale somewhat differently. The scale was standardized on psychopathic deviates, who are characterized by several major adjustment tendencies such as difficulty with authority figures, delinquency, opportunism, and under-inhibition of impulses. In the absence of other information it is difficult to rule out any of the above tendencies as explanations for behavior in prison camp. However, it seems reasonable to suppose that since ***both resisters and collaborators*** were high on *Pd*, it is the under-inhibition of impulses that is common to these two groups.

Those men who were not able to inhibit impulses when under stress may have been forced by their own acting-out tendencies into taking some kind of stand with respect to Chinese indoctrination pressures: they either gave in to them or arbitrarily resisted them.

It will also be recalled that both resisters and collaborators tended to be older, better educated, more intelligent, and more experienced in the army. This combination of factors suggests that "self-confidence" may have been an important element in determining whether a man would take some action: either collaboration or resistance. If this explanation is correct one would also expect the extreme groups to have higher rank. In fact, the resisters and collaborators did tend to have higher rank, but the relationship did not reach statistical significance.

These two factors, acting-out tendencies and self-confidence, suggest that the underlying behavioral continuum to which the various predictors are related is not ***resistance-collaboration*** but ***action-inaction***. Consistently, the chief contrast is between the neutrals, who defended themselves with a wall of apathy (as described psychiatrically in 11), and the two other groups, both of which were characterized by some positive action toward the Communists. This action might take the form of open resistance, of pseudo-collaboration which it was vainly hoped could be used as an undercover means of resistance, or of genuine collaboration regarded as the best way of

dealing with an otherwise hopeless situation. Those men who were low in self-confidence or in acting-out tendencies would be more hesitant to commit themselves either to open resistance or to collaboration, but would tend to follow the passive middle road of the majority.

A recently completed study of former prisoners of war (9) indicates that a higher percentage of resisters and collaborators than of neutrals had been entertainers, were athletically inclined, and were married. If these findings can be interpreted as reflecting tendencies to respond actively rather than passively to situations, they support our interpretation.

The degree to which ideological change was the basis of collaboration is difficult to assess, because none of the tests used here measured such a variable, nor was there evidence concerning a man's political attitudes prior to capture. It should be noted that there is no *necessary* relationship between collaboration and ideological change. It is possible that ideological change preceded and caused collaboration, or that it followed as a consequence of collaboration, or that the two were entirely unrelated, or that both are a function of some other determinant. Furthermore, data from interviews obtained by the senior author (10) suggest that very few, if any, of the men underwent a significant degree of ideological change. Our results are consistent with the interpretation that ideological change was *not* an important determinant of prison camp behavior. Instead, it appears that the tendency of a person to deal with a stressful situation either actively or passively is a more important determinant of his course of action.

SUMMARY

A sample of 759 American army prisoners of war was given a series of biographical questions, intelligence scales, and personality inventories shortly after their repatriation from communist prison camps in Korea. Comparisons were then made among those men who had (*a*) collaborated with the Communists, (*b*) actively resisted them, and (*c*) taken a "neutral" course between these two extremes. Resisters and collaborators were both significantly higher than neutrals in length of internment, length of service, age, intelligence, and the *Pd* scale of the MMPI, but resisters and collaborators did not differ significantly from one another. No significant differences among the groups were found in army rank, civilian occupation, religion, location of home community, or number of parents present in the home. The *A* and *R* scales from the MMPI and a questionnaire designed to measure "authoritarian" attitudes also failed to discriminate. Although differences obtained were too small to be of any practical value, they are of theoretical interest and were interpreted in terms of an activity-passivity hypothesis concerning response to stress.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. HATHAWAY, S. R., & MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory*. (Rev. ed.) New York: The Psychological Corporation, 1951.
3. HUNTER, E. *Brainwashing: the story of the men who defied it*. New York: Farrar, Straus, & Cudahy, 1956.
4. LIFTON, R. A. Home by ship: reaction patterns of American prisoners of war repatriated from North Korea. *Amer. J. Psychiat.*, 1954, **110**, 732–739.
5. MAYER, W. E. Why did many GI captives cave in? *U. S. News and World Report*, 1956, Feb. 24, 56–72.
6. MEERLOO, J. A. M. *The rape of the mind*. Cleveland: World, 1956.
7. PASLEY, VIRGINIA. *21 stayed*. New York: Farrar, Straus, & Cudahy, 1955.
8. SEGAL, H. A. Initial psychiatric findings of recently repatriated prisoners of war. *Amer. J. Psychiat.*, 1954, **111**, 358–363.
9. SEGAL, J. Factors related to the collaborative and resistive behavior of the U. S. Army POWs in Korea. George Washington Univer.: Human Resources Res. Office. (To be published in Human Resources Res. Office Tech. Rep.)
10. SCHEIN, E. H. The Chinese indoctrination program for prisoners of war. *Psychiatry*, 1956, **19**, 149–172.
11. STRASSMAN, H. D., THALER, MARGARET, & SCHEIN, E. H. A prisoner of war syndrome: apathy as a reaction to severe stress. *Amer. J. Psychiat.*, 1956, **112**, 998–1003.
12. WELSH, G. S. A factor study of the Minnesota Multiphasic Personality Inventory using scales with item over-lap eliminated. *Amer. Psychol.* 1956, **7**, 341–342. (Abstract)

Received October 12, 1956.

THE RELATIONSHIP OF AGE AND PRIVILEGE STATUS TO REACTION TIME INDICES OF SCHIZOPHRENIC MOTIVATION[1]

GERALD ROSENBAUM, JAMES L. GRISELL, AND WILLIAM R. MACKAVEY[2]

Wayne State University and Northville State Hospital

A PREVIOUS experiment (9) presented evidence of a schizophrenic deficit in social motivation on a reaction time task. It was found that biological motivation aroused by means of shocks to a reacting finger tended to eliminate the characteristic slowness and variability of schizophrenic performance. Some of the more disorganized schizophrenics, however, were unable to improve significantly even under shock motivation. The presence of this lessened biological reactivity in some of the patients led to the rejection of the postulate of a unitary deficit in the social motivation of schizophrenics and to the consideration of a multiple deficit conception of schizophrenia. Specifically, it was hypothesized that lessened biological reactivity constitutes a dimension of schizophrenic motivation that is relatively independent of the previously reported deficit in social motivation.

The present experiment was designed to explore this hypothesis further by investigating the relationship of age and privilege status to reaction time under shock motivation. If it could be shown that slow reactions to shock are related to age and not to poor socialization, defined as a lack of ground privileges, then defective biological motivation would be partially identified as a relatively independent dimension of schizophrenic deficit. Similarly, the independence of the dimension of defective social motivation would be indicated by a relationship between privilege status and reactions to social stimulation, essentially unaffected by age. Finally, it was hypothesized that young, privileged patients may not exhibit deficit on either of these dimensions, but may be differentiated from normals by displaying defective reactivity to unpredictable stress as a reflection of disabling anxiety motivation.

[1] This study was supported by a grant from the Research Division, Michigan Department of Mental Health.

[2] Now at Michigan State University.

In order to test these hypotheses, schizophrenic subjects (*Ss*), subdivided for age and privilege status, and a group of normal *Ss* were tested on a reaction time task in which they were required to lift a finger from a telegraph key in response to a buzzer. All *Ss* were tested under shock, anxiety, and social motivation conditions. In addition, the Henmon-Nelson, a group paper-and-pencil intelligence test (4), was administered to all *Ss* to obtain some indication of their current mental status. Since Shakow (11) and others (10) have suggested that paper-and-pencil intelligence tests indicate cooperativeness and the ability to sustain attention when used with schizophrenics, the Henmon-Nelson scores were employed as an index of general mental efficiency in assessing the construct validity of the reaction time measures of motivation. The reaction time and Henmon-Nelson tests were later readministered to the patients to determine the reliability of these indices within a schizophrenic population. It was hypothesized that the reaction time scores obtained under each of the motivational conditions would be significantly related to age, privilege status, and the normal-patient classification in the manner indicated.

METHOD

Subjects

Forty female schizophrenic patients from Northville State Hospital were employed as the experimental *Ss*, since females had shown the least improvement with shock motivation in the previous experiment. Two conditions of age and privilege status were arranged in a factorial design by subdividing 20 patients over 40 years of age and 20 under 40 years into two equal groups, each consisting of 10 patients with ground privileges and 10 without privileges. Ten college students from Wayne State University, five men and five women, served as normal controls. Since young students, particularly males, are likely to perform better than schizophrenic females on a reaction time task (9), this control group was employed primarily to obtain an index of optimal functioning against which the schizophrenic groups could be compared. The mean ages in years for the groups were as follows: unprivileged-old 49.7, range 43–56; privileged-old 51.2, range

41–59; unprivileged-young 28.7, range 18–37; privileged-young 28.8, range 23–35; normal controls 20.2, range 17–25.

Apparatus

The reaction time device previously described (9) was modified so that the stimulus presentation, recording, and response units could be housed in a portable wooden cabinet 24½ in. long, 12 in. wide, and 12 in. high. The response unit consisted of a telegraph key and a copper plate mounted on a folding panel hinged to the *S*'s end of the cabinet. When the response panel was opened flat on a table, a 3 x 3 in. opal flash screen mounted on the front face of the cabinet was visible. The illumination of this screen by a 10-w. bulb inside the cabinet served as the ready signal.

The ready signal was presented by experimenter's (*E*'s) closing a push-button type switch. Depression of the telegraph key by *S* in response to the ready signal activated a Hunter Decade Interval Timer and closed a holding relay, permitting *E* to release his switch. At the termination of the interval, previously set on the interval timer by *E*, a door-bell buzzer and Standard Electric Timer were instantaneously activated. Release of the key in response to the buzzer broke the common circuit to the buzzer and the clock. Reaction times were read directly to the nearest .01 sec. The clock, interval timer, and all switches were mounted on the rear face of the cabinet, and were accessible when a folding panel was opened.

The shocks, generated by a Harvard Inductorium located inside the cabinet and wired in parallel with the buzzer and the clock, were administered through a copper electrode on the surface of the response key and the copper plate on which *S* rested his hand. The inductorium was adjusted to transmit a 50-v. current, the setting of a silent mercury switch determining whether or not the shock would accompany the buzzer.

Procedure

All *S*s were instructed that they were taking a reaction time test requiring them to depress the key to the light signal and to release it as rapidly as possible at the sound of the buzzer. The test series consisted of 80 successive trials involving social, anxiety, and shock motivation presented in that order.

Social motivation. The *S*s were required to respond to the buzzer presented at a constant 10-sec. preparatory interval for 15 successive trials, followed by 10 additional trials at this interval during which *S* was told his last reaction time and asked to improve his score. The mean of the first 15 reaction times served the individual social motivation scores.

Anxiety motivation. The *S*s were instructed to be prepared for long and short ready intervals as well as occasional shock, and then presented with 20 trials in which the 1- and 10-sec. intervals occurred equally often in a randomized order. Twenty more irregular trials in which shock was presented on three of the ten 1-sec. trials and three of the ten 10-sec. trials in an unpredictable fashion followed immediately. The mean of the seven reaction times on the 1-sec. nonshock trials served as the anxiety motivation score because it was anticipated that these trials would be most likely to reflect disruptive anxiety in long reaction times.

Shock motivation. The *S*s were informed that they would feel skin stimulation at the same instant the buzzer sounded on all of the last 15 shock trials presented at the regular 10-sec. interval. The mean of these 15 reaction times served as the shock motivation score, to be compared with the 15 trials of the test series given under social motivation.

Form B of the Henmon-Nelson Test of Mental Ability for college students was administered to all patients during the same week that they were tested for reaction time. The patients were tested in small groups of less than 5 persons and aided in understanding the directions in order to maximize their performance on the test. The reaction-time and Henmon-Nelson tests were readministered two to four weeks from the original testing to all patients then available. Of the total sample, five of the privileged-young, two of the unprivileged-young, and one of the unprivileged-old patients were unavailable. Henmon-Nelson scores were available on the normal controls from a classroom administration; no retest was conducted.

RESULTS

The effects of the three conditions of motivation upon reaction time in the schizophrenic and normal groups are presented in Table 1. Since a few extreme cases in the unprivileged-old schizophrenic group tended to influence the group means unduly, medians are also shown in the tables to permit an evaluation of the group trends relatively unaffected by extreme scores. Statistical analyses of the group differences have also taken the heterogeneity of this group into consideration, square root transformations of scores being employed where the group variances were not homogeneous. While the

TABLE 1

REACTION TIMES IN .01 SEC. OF SCHIZOPHRENIC AND NORMAL GROUPS UNDER DIFFERENT CONDITIONS OF MOTIVATION

	Unprivileged Schizophrenics		Privileged Schizophrenics		
	Old	Young	Old	Young	Normals
Shock Motivation					
Mean	19.2	15.1	17.0	13.7	12.0
Mdn.	18.5	14.0	16.5	13.5	12.0
SD	6.2	4.2	4.4	2.3	1.6
Social Motivation					
Mean	146.8	38.3	35.6	23.1	18.8
Mdn.	84.5	34.5	33.0	22.5	18.0
SD	161.1	20.0	16.3	16.8	2.6
Anxiety Motivation					
Mean	124.7	51.5	73.5	32.7	24.8
Mdn.	96.0	45.0	60.0	32.5	24.5
SD	99.8	26.7	41.6	9.4	5.5

normal group was not comparable to the schizophrenic groups in age or sex composition, their average scores are shown in the tables, and comparisons have been made where it was intended to match schizophrenic *S*s against an optimal level of normal functioning.

Shock motivation. The hypothesis that older schizophrenics show a reduced capacity to react rapidly to biological shock motivation appears to be clearly supported by the data shown in the first two rows of Table 1. Analysis of variance of these scores revealed that age was a significant source of variance at the .05 level, while privilege status did not significantly seem to affect reaction time. There was no significant interaction between age and privilege status. A separate *t* test showed that the normals were not significantly faster than the privileged-young patients.

Social motivation. The middle two rows of Table 1 show that the older patients and those without privileges reacted more slowly under social motivation. Analysis of variance indicated that both these sources of variance were significant at the .05 level; their interaction was not statistically significant.

Although these findings are not consistent with the hypothesis concerning the social motivation dimension, the data of both the present and the previous studies showed that the patients who could not react rapidly to shock were also unable to react rapidly to social instructions. Therefore, an index of social motivation was computed for every *S* which takes into account his biological capacity to react as indicated by reaction time under shock.

These social efficiency percentages, based on the ratio between speed under social motivation to speed under shock motivation, are presented in Table 2. The analysis of these data revealed that the privileged patients

TABLE 2
SOCIAL EFFICIENCY PERCENTAGES OF SCHIZOPHRENIC AND NORMAL GROUPS

	Unprivileged Schizophrenics		Privileged Schizophrenics		
	Old	Young	Old	Young	Normals
Mean	31.2	45.6	56.3	60.0	63.0
Mdn.	26.0	45.0	49.0	59.5	63.5
SD	17.5	16.7	27.9	8.6	7.5

TABLE 3
HENMON-NELSON TEST SCORES OF SCHIZOPHRENIC AND NORMAL GROUPS

	Unprivileged Schizophrenics		Privileged Schizophrenics		
	Old	Young	Old	Young	Normals
Mean	7.2	16.6	14.0	27.5	47.1
Mdn.	6.0	18.0	12.5	22.5	46.5
SD	5.3	11.2	10.6	12.6	7.4

obtained significantly higher percentages than the unprivileged patients, but that age did not significantly affect these scores. All the normals and 8 of the 10 privileged-young patients showed greater than 50 per cent social efficiency: these two groups were not significantly different on this measure. These results are in line with the social motivation hypothesis.

Anxiety motivation. In keeping with the previous findings, it may be seen from the last two rows of Table 1 that the groups reacting slowly to shock and social motivation were also slowest under anxiety motivation. Age was a significant source of variance at the .01 level, while privilege status and the interaction were not significant. Furthermore, the privileged-young patients, who reacted at the normal level under shock and social motivation, were significantly slower under anxiety motivation than the normal group (.05 level).

Intelligence test and correlation data. The current mental functioning of the groups is indicated by the Henmon-Nelson scores, shown in Table 3. Since this form of the Henmon-Nelson test was standardized for college students and required concentration in following detailed clerical directions, the scores obtained by the patients might be expected to reflect both the social motivation to cooperate and the capacity to sustain attention. Analyses of these data revealed that youth, privilege status, and membership in the normal group were all significantly associated with higher scores.

The relationship of these mental efficiency scores to the reaction time indices of motivation is shown in the intercorrelation matrix presented as Table 4. The correlations indicate that Henmon-Nelson scores were significantly related to all three measures. Shock and anxiety reaction times were also significantly

related; but social efficiency percentages were not significantly related to either of the other reaction time measures. Furthermore, it may be seen from this table that all these indices had useful retest reliabilities for measurement purposes; shock reaction time and Henmon-Nelson scores being highly stable.

In view of the fact that the Henmon-Nelson scores appeared to provide an index of general mental efficiency significantly related to the other measures, a multiple correlation coefficient was computed to determine the amount of variance in Henmon-Nelson scores that could be accounted for by all three reaction-time scores. This multiple R was equal to .57, accounting for 33 per cent of the variance in Henmon-Nelson scores. The multiple regression equation was:

$$V' = 20.17 - .74X_1 + .24X_2 - .02X_3$$

where

V' = predicted Henmon-Nelson score
X_1 = shock reaction time
X_2 = social efficiency percentage
X_3 = anxiety reaction time

It is apparent from the regression weights and the beta coefficients obtained that most of the variance in mental efficiency in this sample of schizophrenics was attributable to the shock motivation score (12 per cent) and the social efficiency percentage (17 per cent). Anxiety reaction time, apart from that component related to shock reaction time, contributed less than four per cent to these schizophrenic mental efficiency scores, but would probably contribute appreciably to the variance of a distribution composed only of early schizophrenic scores.

TABLE 4

RELIABILITIES AND INTERCORRELATIONS OF REACTION TIME AND HENMON-NELSON SCORES IN SCHIZOPHRENICS*

	Shock Reaction Time	Social Efficiency Percentage	Anxiety Reaction Time	Henmon-Nelson Score
Shock reaction time	.94			
Social efficiency percentage	−.07	.79		
Anxiety reaction time	.64	−.26	.71	
Henmon-Nelson score	−.39	.44	−.39	.91

* The reliabilities presented are Pearson r's based on the scores of the 32 patients available for both the initial test and the retest. The remaining intercorrelations are Pearson r's based on initial test scores of all 40 patients. For 38 df a Pearson r of .31 is required for the .05 level of confidence.

DISCUSSION

The results of the present experiment suggest that three relatively separate dimensions of motivational deficit may be identified among the schizophrenias. There seems to be a class of schizophrenics who show a defective capacity to respond to biological stimulation in addition to the more typical reduction in social motivation. These patients tend to be older and to have poorer mental efficiency but, within the limits of their reduced biological reactivity, they are as socially efficient as other schizophrenics. Such patients may well show the quantitatively subnormal adrenal responses to stress observed by Hoagland (5) in chronic schizophrenics. It would also be profitable to investigate whether such psychotics display a significant number of organic signs on the Rorschach. An expectation of this sort is supported by the finding of Brackbill and Fine (2) that typical process schizophrenics are indistinguishable from organics in the incidence of Piotrowski's ten signs of organic involvement.

A second class of schizophrenics seems unimpaired in biological motivation, but the reactivity of these patients to social incentives is typically well under 50 per cent of their reactivity to shock. The extent of the deficit in social motivation appears unrelated to age but is reflected in a lack of ground privileges and poorer mental efficiency. This group illustrates the social disarticulation process described by Cameron (3) and others (12) in the typical disturbances in reaction time reported originally by the Worcester group (6, 8) and more recently by King (7). King, however, has rejected the motivational interpretation of slowed reaction times, suggesting rather that this symptom is "... a primary indication of disturbance in a basic adaptational process, that of motor adjustment to the external surround." The findings of both the present and previous papers that shock motivation eliminates this primary disturbance in motor adjustment, at least in some patients, makes his interpretation very questionable.

A third class of schizophrenics functions at

a normal level of efficiency under both biological and social motivation. These patients are young, have ground privileges, and tend to display good mental efficiency. They seem to display abnormal reaction times only under the stress of anxiety motivation, suggesting a type of psychosis in which the disorder constitutes primarily a reaction to disruptive anxiety. While the intellectual regression of the first class of schizophrenics appears to resemble Beck's (1) reaction Form S-1, the intellectual control and lack of symptoms of this third class may well involve the defensive character hardening that he has described in his Form S-3.

The present findings do not indicate whether these three types may best be regarded as different forms of schizophrenia or as stages of the same schizophrenic process. However, the reaction-time technique employed appears to provide a brief, reliable instrument for assessing relatively fundamental pathology in motivation. It seems likely that prognosis is poorer in those patients showing a deficit in biological motivation, better in those with good biological motivation and high social efficiency percentages. Further exploration of the manner in which these dimensions of motivation are related to different prognostic and etiological factors should determine whether the types isolated require classification as separate forms of psychopathology or as various phases of a basic schizophrenic process.

SUMMARY

The present experiment was designed to investigate the relationship of age and ground privilege status to reaction time indices of biological, social, and anxiety motivation in schizophrenic patients.

Two conditions of both age and privilege status were arranged in a 2 x 2 factorial design by subdividing 20 schizophrenics over 40 years and 20 under 40 years into two equal groups, each consisting of 10 privilege and 10 ward patients. Ten college students served as normal controls.

All *Ss* were tested on a reaction time task requiring them to lift a finger from a telegraph key in response to a buzzer during 80 successive trials, successively involving social instructions, unpredictable anxiety cues, and shock stimulation. The *Ss* were also given the Henmon-Nelson test of Mental Ability and retested on both the reaction time and intelligence tests.

The results showed that age significantly slowed mean biological and anxiety reaction times but did not affect the social efficiency percentage based on the ratio of speed under social motivation to speed under shock motivation. Privilege status significantly improved mean social reaction time and social efficiency percentage but had no effect on the other two measures and did not interact with age. The normals were superior to the schizophrenics on all measures, but could be significantly differentiated from the younger, privileged group only on anxiety reaction time.

Correlational data were presented showing that shock reaction time was correlated with anxiety reaction time but independent of social efficiency percentage in the schizophrenic sample. Mental efficiency was significantly correlated with the three reaction time indices, and all four measures had stable retest reliabilities.

It was concluded that lessened reactivity to biological stress with increasing age, reduced responsiveness to social demands, and inability to control anxiety are three separate dimensions of motivational deficit which may be isolated in schizophrenic patients. It is suggested that these dimensions may be related to different etiological and prognostic factors, and may prove useful in the classification of schizophrenic pathology.

REFERENCES

1. BECK, S. J. *The six schizophrenias*. New York: American Orthopsychiatric Ass., 1954. Research Monograph No. 6.
2. BRACKBILL, G. A., & FINE, H. J. Schizophrenia and central nervous system pathology. *J. abnorm. soc. Psychol.*, 1956, **52**, 310–313.
3. CAMERON, N. The functional psychoses. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944. Pp. 861–921.
4. HENMON, V. A. C., & NELSON, M. S. *The Henmon-Nelson tests of mental ability*. New York: Houghton-Mifflin, 1931.
5. HOAGLAND, H. Stress and the adrenal cortex with special reference to potassium metabolism. *Psychosom. Med.*, 1950, **12**, 142–148.
6. HUSTON, P. E., SHAKOW, D., & RIGGS, L. A. Studies of motor function in schizophrenia: Reaction time. *J. gen. Psychol.*, 1937, **16**, 39–82.

7. KING, H. E. *Psychomotor aspects of mental disease.* Cambridge: Harvard Univer. Press, 1954.
8. RODNICK, E. H., & SHAKOW, D. Set in the schizophrenic as measured by a composite reaction time index. *Amer. J. Psychiat.*, 1940, **97**, 214–225.
9. ROSENBAUM, G., MACKAVEY, W. R., & GRISELL, J. L. Effects of biological and social motivation on schizophrenic reaction time. *J. abnorm. soc. Psychol.* 1957, **53**, 364–368.
10. SENF, RITA, HUSTON, P. E., & COHEN, B. D. Thinking deficit in schizophrenia and changes with amytal. *J. abnorm. soc. Psychol.*, 1955, **50**, 383–387.
11. SHAKOW, D. The nature of deterioration in schizophrenic conditions. *Nerv. ment. Dis. Monogr.*, 1946, No. 70.
12. SULLIVAN, H. S. *Conceptions of modern psychiatry.* Washington D. C.: W. A. White Foundation, 1947.

Received October 23, 1956.

THE EFFECTS OF SEX-ROLE IDENTIFICATION UPON PROBLEM-SOLVING SKILL[1]

G. A. MILTON

Stanford University

DURING recent years, there has been an increased interest in the effects of "nonintellectual" processes upon problem solving and thinking. One area in which such effects can be and have been studied with readily discriminable criterion groups is the area of sex differences. Sweeney (5) has summarized several studies in which problem-solving differences between men and women were observed. He also reports experiments of his own which demonstrate that men solve certain classes of problems with greater facility than do women, even when differences in intellectual aptitude, academic training, and special abilities are controlled.

The present study explores the possibility that the differences in problem-solving skill between men and women may be due, at least in part, to a set of learned behaviors that constitute a culturally defined sex role. A simple extrapolation from the earlier studies of sex differences leads to the hypothesis that the more an individual identifies with the masculine sex role, the greater will be his problem-solving skill.

This hypothesis may be stated more explicitly: (*a*) There is a positive relationship between masculine identification and problem-solving achievement. (*b*) When an adjustment is made for the between-subjects variance contributed by sex-role identification, sex differences in problem-solving achievement will be reduced. The implication of the hypothesis is, of course, that not only will differences between men and women in problem-solving skill be reduced by an adjustment for masculine role identification, but also that the more "masculine" men will solve problems more readily than "feminine" men and, similarly, that the more masculine women will be better problem solvers than women who have greater identification with the feminine role.

[1] This experiment is one in a series of studies of problem-solving done under Project NR 150-149 and supported by Contract N6onr 25125 between Stanford University and the Office of Naval Research. Work on the contract was under the general direction of Dr. Donald W. Taylor. Permission is granted for reproduction, translation, publication, use, and disposal of this article in whole or in part by or for the United States Government.

METHOD

Indices of Sex-Role Identification

The concept of sex role refers to a heterogeneous collection of actions that differentiate typically masculine behavior from typically feminine behavior in a particular culture. These characteristics of the sex role may be defined restrictively in terms of overt behavior only, or more broadly in terms of attitudes, interests, needs, and other inferred constructs in addition to overt behavior. The indices of sex role used in this investigation have been determined by primarily empirical means; overt and covert measures were used, the criterion for selection being that they involved nonintellectual factors that differentiate men from women for the general culture from which the experimental sample was drawn.

The Terman-Miles M-F Test (6) was chosen as the primary defining measure of sex-role identification because it measures a broader range of responses than any other sex-role measure cited in the literature. This test consists of a set of items that differentiate men from women in a culture similar to that of the current sample (cf. below). The test elicits responses that not only discriminate men from women, but also seem to assess a culturally defined masculinity-femininity variable within each sex.

Two secondary measures of sex role were also employed. These tests tap a narrower range of sex-role content. One, the M-F scale of the *Minnesota Multiphasic Personality Inventory* (2), consists largely of self-report items referring to feelings and emotions. The other, the *Behavioral Inventory* M-F (4) scale, is part of an unpublished experimental inventory that defines the role in terms of self-reported overt behavior that discriminates between men and women in a college population.

Problems

Twenty problems that had shown significant sex differences in pretest trials were used. The problems were selected and refined by Nakamura (3), and his methods of administration were duplicated in the present study. Item analysis has indicated that the problems involve two types of problem-solving skill, restructuring and straightforward solution. The restructuring problems require the subject to alter his initial set in order to attain the solution; the straightforward problems are solvable by direct means. Ten straightforward and ten restructuring problems were arranged alternately by pairs, according to order of difficulty. One member of each pair involved a numerical solution and the other member was non-numerical.

The problems were bound in a booklet with one problem per page so that each problem was presented individually. A working time of four minutes per problem was allowed with a ten-minute intermission after the first ten problems. Problems were scored either correct or incorrect with no partial credits given.

Subjects

One hundred and twenty-nine students from an introductory psychology class at Stanford participated as subjects. Of these, 63 were males and 66 were females. Scores on the mathematical and verbal sections of the College Entrance Examination Board Scholastic Aptitude Test were available for all subjects as a control upon intelligence.

Procedure

Subjects were tested in six groups of 16 to 35 subjects. For all groups, the tests were administered in three sessions. The order of presentation of the tests was arbitrary. The MMPI M-F and the Behavioral scale M-F were presented during the first session; the problem-solving test during the second session, and the Terman-Miles M-F during the third session. Because of the nature of the items on the M-F scales, it was assumed that there would be no interaction with order of presentation; therefore, this arrangement was made purely on the basis of experimental expediency.

RESULTS

The results of this investigation may be divided into two categories. First, sex differences on the problem-solving measure are examined in order to establish that those reported by other investigators (5) are found in the current investigation. The hypothesized relationship between sex role and problem solving is then tested. Both the relationships between sexes and within sexes are examined.

Sex Differences in Problem Solving

The differences between the problem-solving achievement of men and women presented in Table 1 are the data which the hypothesis must partially explain. The differences between men and women are significant with respect to the total problem-solving score and all the sub-scores, but the difference between men and women on one subtest is not significantly greater than that on another subtest. The total score can therefore be used to test the effects of the sex-role measure. These results are similar to those obtained by Nakamura (3) for the same problem on a sample of University of California students.

TABLE 1
SEX DIFFERENCES IN PROBLEM-SOLVING SKILL

Type of Problem	Men		Women		t
	Mean	*SD*	Mean	*SD*	
Problem-solving, total score	9.10	3.68	7.18	3.27	3.12**
Restructuring problems	4.33	2.22	3.38	1.88	2.63*
Straight-forward problems	4.76	2.17	3.80	1.79	2.74**
Numerical problems	5.06	2.27	4.04	1.87	2.78**
Non-numerical problems	4.05	2.15	3.14	1.79	2.59*

* Significant beyond .02 level.
** Significant beyond .01 level.

TABLE 2
CORRELATION OF THE SEX-ROLE AND APTITUDE VARIABLES WITH PROBLEM-SOLVING SKILL

Variable	Men	Women	Combined
Terman-Miles M-F	.44**	.26*	.41**
MMPI M-F	.14	.13	.29*
Behavior M-F	.28*	.23	.36**
Verbal aptitude	.44**	.28*	
Math aptitude	.63**	.78**	

* Significant beyond .05 level.
** Significant beyond .01 level.

Effects of Sex-Role upon Sex Differences in Problem Solving

In order to confirm the hypothesis of this investigation, two effects must be demonstrated: that there is a correlation between sex-role identification and problem solving, and that this correlation accounts for some of the differences between men and women in problem-solving skill. The correlations between sex-role measures and problem solving are presented in Table 2.

It is clear that problem-solving skill is significantly related to sex-role identification as measured by the Terman-Miles test. There is a significant relationship both within and across sexes. The evidence concerning the secondary measures of sex-role is less clear. The correlation of problem solving with the MMPI M-F scale approaches significance only for the combined sex group, and the correlation of problem solving with the Behavioral M-F scale does not appear to be significant for the women.

In order to determine if the differences between men and women in problem-solving score can be accounted for in terms of score on the measures of sex-role identification, a simple analysis of covariance was performed. The rationale of this operation is that if the

TABLE 3

ANALYSIS OF COVARIANCE: SEX DIFFERENCES IN PROBLEM SOLVING WITH ADJUSTMENT FOR TERMAN-MILES M-F SCORE

Source of Variance	Before Adjustment		After Adjustment			
	df	Sum of Squares	*df*	Sum of Squares	Mean Squares	*F*
Total	128	1675	127	1375		
Within	127	1557	126	1366	10.84	
Between	1	118	1	9.32	9.32	0.86

TABLE 4

ANALYSIS OF COVARIANCE: SEX DIFFERENCES IN PROBLEM SOLVING WITH ADJUSTMENT FOR MMPI M-F SCORE

Source of Variation	Before Adjustment		After adjustment			
	df	Sum of Squares	*df*	Sum of Squares	Mean Squares	*F*
Total	128	1675	127	1534		
Within	127	1557	126	1529	12.13	
Between	1	118	1	5.86	5.86	0.48

TABLE 5

ANALYSIS OF COVARIANCE: SEX DIFFERENCES IN PROBLEM SOLVING WITH ADJUSTMENT FOR BEHAVIORAL M-F SCORE

	Before Adjustment		After Adjustment			
	df	Sum of Squares	*df*	Sum of Squares	Mean Squares	*F*
Total	128	1675	127	1462		
Within	127	1557	126	1454	11.54	
Between	1	118	1	7.76	7.76	0.67

correlation between the problem-solving score and the sex-role measure is sufficient to reduce significantly the difference in problem-solving score, then sex-role identification as measured by the particular test accounts for a part of this difference. Table 3 presents the results of this analysis for the Terman-Miles test, and Tables 4 and 5 show similar results for the MMPI and the Behavioral Inventory.

The results indicate that the Terman-Miles not only accounts for a significant part of the difference between men and women in problem-solving skill, but also that it diminishes this difference to the point where it is no longer significant. The results with the MMPI and Behavioral Inventory are comparable to those obtained with the Terman-Miles. The difference between men and women on the problem-solving measure diminishes when the variance attributable to sex-role identification is controlled.

TABLE 6

BETA WEIGHTS: RELATIVE INDEPENDENT CONTRIBUTION OF EACH OF THE VARIABLES TO PROBLEM-SOLVING SCORE

Variable	Men ($N = 63$) Beta	Women ($N = 66$) Beta
Terman-Miles M-F	.297**	.183*
MMPI M-F	.074	.050
Behavior M-F	.066	.126
Verbal aptitude	.224**	.099
Math aptitude	.473**	.848**

* Significant beyond .05 level.
** Significant beyond .01 level.

The hypothesis requires, by implication, that differences in problem-solving skill within a sex must also be partially accounted for by differences in sex-role identification. Table 2 indicates that the Terman-Miles measure was significantly correlated with problem solving within both sexes, but that the other sex-role measures were less clearly related to problem solving.

In order to determine what portion of the variance in the problem-solving task within a sex may be explained on the basis of sex-role identification and the other nonintellectual variables related to sex role, the variables were fitted to a multiple regression equation and the appropriate beta weights were determined for each variable by means of a Doolittle solution. Since the hypothesis requires that differences in problem-solving skill must be explained independently of differences in intellectual aptitude, the verbal and mathematical measures were included in the equation. The results of this analysis are presented in Table 6.

For both the men and the women, the Terman-Miles test measure of sex-role identification is a significant contributor to the variance in problem-solving scores, even when mathematical and verbal aptitude are held constant. The mathematical aptitude scale contributes a large portion of the variance for both sexes, and the verbal aptitude scale makes a significant contribution for the males. The other sex-role variables make no independent contribution to the variance in problem solving, when the more powerful Terman-Miles test is included in the battery.

DISCUSSION

The results of this study seem to indicate that the differences between men and women in problem-solving skill as reported by a number of investigators are merely part of a more general difference between men and women in role identification. Men have characteristically different behavior, attitudes, emotions, and motivations than do women, and the differences in problem-solving skill reflect or are a part of these differences. To the extent that allowance can be made for these differences in general sex-role identification, the differences in problem-solving achievement seem to disappear.

Moreover, there are differences *within* a sex as to the extent of identification with the appropriate sex role, and these within-sex differences in role also seem to be related to problem-solving skill. Men who identify with the masculine sex role are better problem solvers than are those who identify with a more feminine role, and conversely, women who identify with the appropriate sex role have poorer problem-solving scores than do women who have a more masculine identification. All these comparisons, of course, assume a common level of intelligence.

This interpretation leads to several implications. For example, the acquisition of problem-solving skills may be affected by the same processes which govern the formation of sex-role identification. If the Freudian hypothesis concerning the identification process is correct, then important antecedents of adult problem-solving skill are to be found in childhood. The male child who forms an adequate identification with the male role, given sufficient intelligence and opportunity to learn, will grow up to be a good problem solver. On the other hand, the female child, even though possessing adequate intelligence and opportunity to learn, will probably not develop problem-solving skills (of the sort used in this study) if she forms an appropriate identification with the feminine role, because this type of problem solving is not appropriate to the female sex-role in her culture. This is, of course, a very broad extrapolation from the data of this experiment.

Another interesting implication concerns the possibility for improvement in problem-solving skill. If the sex-role interpretation of the results of this experiment is correct, any attempt to change problem-solving skill will be more effective if corresponding changes are made in certain aspects of the conventional sex-role identification. While it may not be necessary to change the sex-role identification of a subject in order to change his (or her) problem-solving skill, it would seem necessary to neutralize the effects of sex-role identification that relate to the problem-solving process. Practice on problems alone probably would not be effective. Carey (1) has demonstrated an effect similar to that of partially changing the problem-solving attitudes associated with sex-role identification. Through group discussions of women's attitudes toward intellectual activity, she was able to improve their problem-solving scores. This change, though significant, was slight, and it would seem that more than a few discussions are required to change attitudes that appear to be basic to the culturally defined sex role. If one considers that sex-role identification has had an effect on the learning of a subject for the major part of a lifetime, it seems that a change in sex role would require a long period of relearning. Also, if major changes in problem-solving performance are to be achieved, the relearning of sex-role would necessarily be accompanied by a period of relearning of problem-solving skills.

Although this interpretation of the results seems appropriate for a college population, there are logical difficulties in any extrapolation of these results to the more general population. Wherever there are subsections of the general population in which problem solving, as defined in this investigation, is not appropriate to the masculine role, this relationship would not be expected. Since the studies of sex differences in problem solving which were mentioned earlier have been conducted upon college populations, there is little evidence as to how far these conclusions can be generalized.

SUMMARY

The hypothesis of the study was that sex differences in problem-solving skill may be partially accounted for by differences in sex-role identification. The Terman-Miles M-F Test was used as the primary index of sex-role identification. Two other M-F ques-

tionnaires were also employed. These scales and a measure of problem-solving skill were administered to 63 male and 66 female university undergraduates. In general, the results indicate that there is a positive relationship between masculine sex-role identification and problem-solving skill both across sexes and within a sex. When allowance is made for this relationship, the difference between men and women in problem-solving performance is diminished.

REFERENCES

1. CAREY, G. L. Reduction of sex differences in problem solving by improvement of attitude through group discussion. Dept. Psychol., Stanford Univer., 1955, (ONR Tech. Rep. No. 9, Contract N6onr 25125).
2. HATHAWAY, S. R., & MCKINLEY, J. C. *Minnesota Multiphasic Personality Inventory*. New York: Psychol. Corp., 1951.
3. NAKAMURA, C. Y. The relation between conformity and problem solving. Dept. Psychol., Stanford Univer., 1955. (ONR Tech. Rep. No. 11, Contract N6onr 25125).
4. PINNEAU, S. R., & MILTON, G. A. The objective measurement of overt behavior. Paper read at Western Psychol. Ass., June, 1954.
5. SWEENEY, E. J. Sex differences in problem solving. Dept. Psychol., Stanford Univer., 1953, (ONR Tech. Rep. No. 1 Contract N6onr 25125).
6. TERMAN, L. M., & MILES, C. C. *Sex and personality*. New York: McGraw-Hill, 1936.

Received October 24, 1956.

THE EFFECTS OF CLEAR AND UNCLEAR ROLE EXPECTATIONS ON GROUP PRODUCTIVITY AND DEFENSIVENESS[1]

EWART E. SMITH[2]

University of Colorado

IN ALL groups, people invest considerable time and energy trying to develop sufficient understanding of other members in order to feel they can predict how these other members will behave. In role theory concepts (5), this is attempting to locate the position of the other so that one may have valid role expectations of the other's role enactments. The same phenomena have been discussed in other theoretical terms. Bales (1) states that the social structure of groups develops in order to "... reduce the tensions growing out of uncertainty and unpredictability in the actions of others."

The general hypothesis of the present study is that the time and energy used by group members in attempting to predict the behavior of others reduces the amount of group energy available for any given group goal or task. This hypothesis is consistent with the conception of group energy formulated by Cattell (2), who states that group members have a finite amount of energy available in a group situation, and that the amount of energy absorbed by the internal group structure is subtracted from the amount available for group productivity.

The present investigation is primarily concerned with the relationship between the degree to which members of a small group feel they can predict each other's behavior, and the group's task efficiency. In addition to the positive relationship anticipated between these variables, it was expected that uncertainty regarding the roles of the other group members would be reflected in increased defensiveness and dissatisfaction with the group experience.

While these problems have not previously been studied experimentally, as far as the author can ascertain, Torrance (8) developed similar views from his study of 1,000 critical incidents of Air Force crews downed in enemy territory. He concluded that the two major detrimental factors in group survival were an unclear situation and an unclear group structure. Regarding the latter factor, he states, in part:

> Under stress, these linkages between members may become confused and thus people do not have a clear perception of what they can expect from one another, with whom they can relate, how they can relate to one another, and so on. The fact that this happens apparently influences the ability of the group to survive.

In order to induce ambiguous role expectations, paid accomplices were assigned roles as silent members in small groups of persons who were, initially, strangers to each other. In the initial meeting of such a group, each member should find the behavior of all of the others equally unpredictable. But in the course of group interaction the behavior of those who spoke should become more predictable, while the behavior of those who remained silent should remain unpredictable. Eventually the group members would undoubtedly come to predict with confidence the simple fact that the silent members would remain silent. Since only short-term groups were studied, however, the latter effect is not relevant to the present experiment.

This reasoning was substantiated in several preliminary investigations. The presence of silent group members in short-term groups had little, if any, effect on productivity during the first few minutes, but productivity decreased to a low point at about thirty minutes, followed by a gain until the presence of silent members resulted in no decrement in productivity.

The hypotheses tested in the present study are:

1. The presence of silent members in a group produces (*a*) decreased productivity, (*b*) increased defensiveness, and (*c*) increased dissatisfaction with the group experience.

[1] Condensed from a dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Psychology, University of Colorado. The writer wishes to express his appreciation to Professor Jack R. Gibb, under whom the work was done, and to Professors Maurice P. Smith and Michael Wertheimer for their helpful criticisms and advice. The work was supported by a research grant (Contract Nonr-1147(03), NR 170-226) from the Group Psychology Branch of the Office of Naval Research.

[2] Now with the RAND Corporation.

2. Decreasing the ambiguity of role expectations in a group that contains silent members counteracts the effects predicted in Hypothesis 1.

METHOD

Subjects and Task

The subjects (*Ss*) were 50 males and 90 females from the introductory psychology classes at the University of Colorado. The task, taken from an experiment by Taylor and Faust (7), is an adaptation of the parlor game, "Twenty Questions," in which the *Ss* had to identify items designated by the experimenter (*E*) as either animal, vegetable, or mineral. Each problem had to be solved with less than forty questions for the group to receive credit. There was no time limit on individual problems. To each question asked by a group member, *E* replied in one of the following ways:

(*a*) Yes; (*b*) No; (*c*) Partly; (*d*) Sometimes; (*e*) Not in the usual sense of the word; (*f*) I don't know (no charge for the question); (*g*) Please restate the question (if the question was unclear or could not be answered in one of the above ways). Difficult items like wrench, ruby, and bread were used.

Experimental Design

In the experimental Conditions A and B, each group was composed of three participating *Ss*, two females and one male, and two silent members, one female and one male. Condition A was designed to induce ambiguous role expectations. Condition B differed from Condition A in that the roles of the silent members were clarified prior to experimentation; all group members checked their "usual role" in a group on a check list. Everyone then announced the role he had checked, which was then checked by *E* on a large replica of the "Usual Roles" check list. The silent members always checked the "Listener" role.

In the control Conditions C and D, there were no paid silent members. In Condition C each group was composed of three participating *Ss*, two females and one male. In Condition D, each group was composed of five participating members, three females and two males.

In all conditions, *Ss* worked on the task for two periods. The first work period, which lasted fifteen minutes, was intended to allow ambiguous role expectations of the silent members to develop. In the second work period, which was ten minutes long, group productivity was measured. The problems, and their order, were the same for all groups. Regardless of how many problems the group had solved in the first work period, each group started with a new, and identical, set of problems in the second work period.

The thirteen silent paid participants were not known to the *Ss* in their groups. Each silent member was assigned to Condition A and B groups an equal number of times.

Dependent Variable Measurements

Group productivity was measured by the number of problems solved. ***Defensiveness*** was measured by a 25-item Likert scale, with a corrected, split-half reliability of .79, which had been validated by the known group method (4). A typical item is "I felt free to express all my ideas in this group." ***Group satisfaction*** was measured by a 15-item Likert scale, with a corrected split-half reliability of .93, which had been validated by the known group method. A typical item is "This is the most stimulating experience I've had in a long time." A single item was used as an index of ***role ambiguity***: "I can predict how this person would act in other situations." The *Ss* rated each group member on a scale from −3 to +3. A numbering system was used so that *Ss* did not give their names. This measure was used as a check on the effectiveness of the experimental manipulation.

TABLE 1

MEDIAN RATING OF *Ss* AND SILENT PAID PARTICIPANTS ON ROLE AMBIGUITY

Statistics on Role Ambiguity Rating	Data under each condition					
	A Ambiguity (Silent Members)		B Clarity (Silent Members)		C Control ($N = 3$)	D Control ($N = 5$)
	Rating of *Ss*	Rating of Silent Members	Rating of *Ss*	Rating of Silent Members	Rating of *Ss*	Rating of *Ss*
Median	11.00	6.00	11.00	8.00	8.50	8.00
N	29	29	30	30	30	50

Note.—Low scores indicate role ambiguity.

H for these data is 34.89, $p < .01$.

N is 29 in Condition A as one *S* failed to fill out the rating form.

RESULTS AND DISCUSSION[2]

Table 1 contains the findings on the ambiguity item, "I can predict how this person would act in other situations." Plotting the scores on this item revealed that normal distributions could not be assumed. Consequently, nonparametric rank tests have been used (3, 6). An H test, which is functionally equivalent to a one-way analysis of variance, has been computed on the data in Table 1. As this H is statistically significant, additional rank tests have been computed. Using this method, z's are obtained which are normal deviates evaluated by the use of the normal probability table. The z (for related measures) on the ratings of silent members compared to the ratings of *Ss* in the Condition A groups is 4.46**, indicating that the independent variable manipulation in Condition A was successful; i.e., that the silent paid participants were seen as less predictable than the *Ss*.

[2] Statistical significance at the .05 level has been indicated by a single asterisk; at the .01 level or better by a double asterisk.

Additional support for the hypothesis that silent members were seen as less predictable is given by the significant z, 4.37**, between the ratings of silent members and *Ss* in the Condition B groups. However, comparison of the ratings of the silent members in Condition A with the ratings of the silent members in Condition B indicates that the difference, although in the predicted direction, is not significant. Therefore, the hypothesis that role clarification reduces the ambiguity of the roles of silent members is not supported.

This failure may be due to a temporal factor. At the time that the *Ss* rated the silent members on predictability, they had a period of 30 minutes of observations upon which to base their ratings. This 30-min. time sample was the same in Conditions A and B, and was apparently more powerful than the earlier one-minute role clarification in Condition B, perhaps resulting in a leveling effect on the comparison of Conditions A and B with respect to the ratings of the silent members. Results more in line with the hypothesis might have been obtained if the rating item had been given earlier.

An interesting, and unpredicted, finding is that the mean rating of the *Ss* in both Conditions A and B is significantly higher than the ratings of *Ss* in Condition C, with z's of 3.24** and 3.64**, using a rank test for independent samples (3). This finding cannot be explained from the data, but, as a possible explanation, it may be suggested that in the groups containing silent members, the other *Ss* were seen by contrast as much more predictable than they otherwise would have been.

Table 2 contains the findings on group productivity. Plotting the group task scores for the second work period (as anticipated, there were no significant differences in the first work period) revealed that normal distributions could not be assumed. Consequently, nonparametric rank tests (for independent measures) were used. The obtained z between Conditions A and B is 2.18*; between A and C, 2.35*; and between A and D, 2.63**.

Table 2 shows that the silent members reduced group productivity significantly in Condition A. These results support the hypothesis that ambiguous role expectations reduce group productivity. At first glance, some more parsimonious explanations might be suggested. Thus, one might suppose that the decrement in productivity in Condition A was produced by silent members per se, rather than by ambiguous role expectations. However, silent members were also present in the Condition B groups, where they were as silent, "uncooperative," "unproductive," etc., as they were in the Condition A groups. The difference was that in the Condition B groups, the *Ss* expected the silent members to be silent. Therefore, the variable to which the decrement in productivity must be attributed was the unpredictability of the silent members.

TABLE 2

NUMBER OF PROBLEMS SOLVED, UNDER EACH OF FOUR CONDITIONS, IN FIRST AND SECOND WORK PERIODS

Statistic on Number of Problems Solved	Data Under each Condition			
	A Ambiguity (Silent Members)	B Clarity (Silent Members)	C Control ($N = 3$)	D Control ($N = 5$)
Median: First work period (15 minutes)	5.50	6.50	6.00	6.00
Median: Second work work period (10 minutes)	0.	3.00	2.00	2.50
Number of groups	10	10	10	10

Note.—H for these data (second work period) is 9.29, $p < .05$.

TABLE 3

AVERAGE DEFENSIVENESS SCALE SCORES UNDER EACH OF FOUR CONDITIONS

Statistic on D-Scale Scores	Data Under each Condition			
	A Ambiguity (Silent Members)	B Clarity (Silent Members)	C Control ($N = 3$)	D Control ($N = 5$)
Mean	66.00	61.53	49.17	52.74
SD	13.92	12.41	14.14	14.10
N	30	30	30	50

Note.—High scores indicate defensiveness.

The analysis of variance is not summarized, to conserve space, as only predicted differences have been tested. However, the overall F on the means is 9.72, $p < .01$.

The data in Table 3 support the first hypothesis, that the presence of silent members produces increased defensiveness. The t ratios between Conditions A and C, and A and D, are 4.48** and 4.04**, respectively. The second hypothesis, that the presence of silent members does not produce increased defensiveness if the roles of the group members are explicit, is not supported. The differences

TABLE 4
AVERAGE GROUP SATISFACTION SCALE SCORES UNDER EACH OF FOUR CONDITIONS

Statistic on GS Scale Scores	Data Under each Condition			
	A Ambiguity (Silent Members)	B Clarity (Silent Members)	C Control ($N = 3$)	D Control ($N = 5$)
Mean	54.80	62.60	74.77	74.50
SD	9.74	13.82	14.81	13.06
N	30	30	30	50

Note.—High scores indicate satisfaction.
The *F* ratio between the variances of Conditions A and C is 2.31; the over-all *F* on the means is 18.34, $p < .01$ in both instances.

between Condition B and Conditions C and D on defensiveness are significant, with *t*'s of 3.45** and 2.79**, respectively. The difference between Conditions A and B is not significant.

The data in Table 4 clearly support the first hypothesis, that the presence of silent members reduces group satisfaction. The *t* between Conditions A and D is 9.26**, and that between Conditions A and C is 6.05**. As the *F* ratio on the variances of Conditions A and C is 2.31*, a correction for *t*'s computed when variances are heterogenous has been used (3). The second hypothesis, that the presence of silent members does not reduce group satisfaction if the roles of the group members are explicit, is partially supported by the data. Group satisfaction is higher in Condition B than in Condition A, with a *t* of 2.48*. However, the decrement in satisfaction produced by the silent members was not completely offset by the role clarification in Condition B groups, as the *t* ratios between Condition B, and Conditions C and D, on group satisfaction, are 3.23** and 3.81**, respectively.

At this point, it should be noted that although the primary independent variable, the presence of silent members, produced the expected results in all cases, the lack of significant differences between Conditions A and B on the role ambiguity item and on the Defensiveness scale throws some doubt on the interpretation of the results given here. One must decide whether Conditions A and B were really different, as indicated by the significant differences between them on productivity and the Group Satisfaction scale; or whether they are really similar, as indicated by the lack of significant differences on the ambiguity item and Defensiveness scale.

The argument here is that Conditions A and B were different, and that role ambiguity and role clarification were the variables which lowered and raised group productivity. The lack of significant differences between Conditions A and B on the role ambiguity item and the Defensiveness scale are interpreted as a product of their relatively low reliability and of the decay of the role clarification set-variable by the time the questionnaires were given. However, the data are open to alternate interpretation.

SUMMARY

In a laboratory investigation, paid participants, who were instructed to remain silent, were included in small groups to induce ambiguous role expectations. Five-member problem-solving groups, each including three naive *S*s and two silent members, were used in each experimental condition. Three- and five-member groups were used in control conditions. It was hypothesized that ambiguous role expectations would reduce group productivity and satisfaction, and increase defensiveness.

All these predictions were well supported by the data.

Silent paid participants were included in another condition, where group members revealed their intended role behaviors. Group members thus had accurate expectations that the paid participants would remain silent. It was hypothesized that, as a result of this role clarification, the effects predicted above would be reduced or negated. Thus, the presence of silent participants would not reduce group productivity and satisfaction, or increase defensiveness.

Role clarification, with silent paid participants present in the groups, did, as predicted, result in normal (i.e., control group level) group productivity. Group satisfaction was significantly increased, but not to the normal level. The prediction that role clarification would restore the decrement in defensiveness was not supported.

REFERENCES

1. BALES, R. F. A theoretical framework for interaction process analysis. In D. Cartwright and

A. F. Zander (Eds.), *Group dynamics: Research and theory*. Evanston, Ill.: Row, Peterson, 1953.

2. CATTELL, R. B. New concepts for measuring leadership, in terms of group syntality. *Hum. Relat.*, 1951, **4**, 161–184.
3. EDWARDS, A. L. *Statistical methods for the behavioral sciences*. New York: Rhinehart & Co., 1954.
4. GIBB, J. R. *Factors producing defensive behavior within groups*. Quarterly Status Report, Office of Naval Res., Contract Nonr-1147(03), NR 170-226, Feb. 15, 1956, 5 pp.
5. SARBIN, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge, Mass.: Addison-Wesley, 1954.
6. SIEGEL, S. *Non-parametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
7. TAYLOR, D. W., & FAUST, W. L. Twenty questions: Efficiency in problem solving as a function of size of group. *J. exp. Psychol.*, 1952, **44**, 360–368.
8. TORRANCE, E. P. The behavior of small groups under the stress of conditions of "survival." *Amer. sociol. Rev.*, 1954, **19**, 751–755.

Received October 24, 1956.

THE RELATIONSHIP BETWEEN OVERT AND FANTASY AGGRESSION AS A FUNCTION OF MATERNAL RESPONSE TO AGGRESSION

GERALD S. LESSER

Hunter College

IN RECENT years, a voluminous literature has developed around the problem of establishing relationships between fantasy behavior and overt behavior. Different researchers have used different drive areas, different populations, different theoretical bases, and different methods of measurement. The most conspicuous conclusion is that the empirical findings are not in agreement.

The importance of this area of investigation for both clinical practice and personality theory has been elaborated by Lindzey (12). He concludes that one of the most important and difficult problems is the "determination of the conditions under which inferences based upon projective material directly relate to overt behavior and the conditions for the reverse" (12, p. 18). The present study concerns the differential conditions under which aggressive behavior is learned that may allow prediction of how aggressive expressions in fantasy are related to those in overt behavior.

Various studies (9, 13, 16, 18, 19) have demonstrated that the degree of correspondence between fantasy behavior and the associated overt behavior is greater for certain drives than for others. Significant positive correlations have been reported between TAT fantasy and overt behavior for variables such as abasement, achievement, creation, dependence, exposition, nurturance, etc. Significant negative correlations have been reported for sex, and inconclusive results have been obtained for a wide variety of other variables. For the variable of aggression, results include significant positive correlations between fantasy and overt expressions (8, 14), significant negative correlations (5, 16), and inconclusive findings (1, 2, 3, 4, 6, 10, 13, 15, 16, 18).

To resolve these inconsistent results, it has been suggested (13, 14, 16, 19) that motives that are culturally encouraged are "... likely to be as strong in their overt as in their covert manifestations" (13, p. 16), while motives that are culturally discouraged are apt to show little or no relationship between the strength of fantasy and overt expressions.

Mussen and Naylor (14) have attempted to test the first segment of this formulation. They contended that lower-class culture encourages aggression, and predicted that "... in a lower-class group, individuals who give evidence of a great deal of fantasy aggression will also manifest more overt aggression than those who show little aggression in their fantasies" (p. 235). A mixed group of white and Negro boys, "... almost all of whom had been referred to the Bureau of Juvenile Research for behaviors which brought them into conflict with school and court authorities..." (p. 236), were used as subjects. The authors report a statistically significant but not especially strong positive relationship between ratings of overt aggression and number of aggressive TAT themes. Further investigation of Mussen and Naylor's hypothesis would profit from more precise measurement of parental response to aggression, control comparisons, and a more representative sample.

The present study seeks to examine the comparative consequences of both encouragement and discouragement of aggression through the hypothesis that under conditions of maternal encouragement of aggression a greater degree of correspondence exists between fantasy and overt aggression of children than under conditions of maternal discouragement of aggression.

METHOD

Subjects

The subjects (*Ss*) were 44 white boys (ages 10-0 to 13-2) and their mothers. The boys were drawn from one fifth grade and two sixth grades in two public schools. All of the boys and their mothers in these three classes participated except one mother who refused to be interviewed. The Kuhlmann-Anderson intelligence quotients of the boys ranged from 82 to 119, with a mean of 102. The two schools are in adjacent districts and the families constitute a relatively homogeneous upper lower-class group.

Maternal Attitudes and Practices

Only one aspect of the environmental conditions of learning of aggressive behavior was measured, i.e., the maternal attitudes and practices supporting or prohibiting aggression. A structured questionnaire-interview schedule was orally administered to the mothers in their homes by a male interviewer. Questions regarding the support or prohibition of aggression constituted only one segment of the total interview; the entire interview schedule is described in detail elsewhere (11). Pertinent to the present study were eight items concerning the mother's attitudes toward aggression in children, and thirteen items about the mother's practices in dealing with the aggressive behavior of her child. An illustrative item measuring maternal attitudes toward aggression is: "A child should be taught to stand up and fight for his rights in his contacts with other children." The four response alternatives of agree, mildly agree, mildly disagree, and disagree were allowed for this item. An example of an item measuring maternal practices concerning aggression is: "If your son comes to tell you that he is being picked on by a bully at the playground who is his own age and size, there would be a number of different things you might tell him. Would you tell him to ignore him and turn the other cheek?" Response alternatives for this item were yes and no. Items that did not involve judgments on a four-point scale were transformed to have approximately the same range of scores as the items that involved four alternatives.

A single score was obtained for each mother by combining all items, assigning plus scores to the responses indicating support of aggression and minus scores to responses indicating discouragement of aggression. The range of scores was from +9 to −7, with a median score of +2. The corrected odd-even reliability coefficient was .80.

The distribution of scores for maternal response to aggression was dichotomized to form one group of mothers (with scores above or at the median) whose attitudes and practices were more supportive of aggressive behavior than those of the other group (with scores below the median). The hypothesis demands that the correlation between fantasy and overt aggression for the children of the mothers in the former group be significantly more positive than the corresponding correlation for the children of the mothers in the latter group.

Fantasy Aggression

Fantasy aggression in the children was measured through an adaptation of the TAT procedure (13, p. 3–5). A set of ten pictures was designed. In each picture two boys are interacting. The pictures differed from one another in the degree to which the instigation to aggression was apparent.

To insure complete and accurate transcription of the stories, tape recordings were taken. An introductory period preceding the fantasy task served both to establish rapport between the child and the male examiner, and to familiarize the child with the recording device. Instructions were:

> I'm going to show you some pictures. These are pictures of two boys doing different things. What I'd like you to do is make up a story to each of these pictures. You can make up any story you wish; there are no right or wrong stories. Say what the boys are thinking and feeling and how the story will turn out.

The ten pictures, in the order of presentation, were:

1. One boy is holding a basketball and the other boy is approaching him with arms outstretched.
2. One boy is stamping upon an ambiguous object and the other boy is reaching for the object.
3. One boy is sitting behind the other boy in a classroom and is leaning toward him.
4. One boy is walking down the street and the other boy, with fists clenched, is glaring at him.
5. One boy, with fists clenched, is staring at the other boy who is sitting, head bowed, on a box.
6. One boy is sawing a piece of wood and the other boy is leaning on a fence between them, talking to him.
7. The two boys, surrounded by a group of other boys, are approaching each other with arms upraised and fists clenched.
8. The two boys are making a fire. One boy is kneeling to arrange the wood and the other boy is approaching, ladened with wood for the fire.
9. One boy, who is looking back, is running down a street and the other boy is running behind him.
10. Two boys are standing in a field. One boy, with his hand on the other boy's shoulder, is pointing off in the distance.

A fantasy aggression score was obtained for each *S* by counting the number of times the following acts appeared in his stories: fighting, injuring, killing, attacking, assaulting, torturing, bullying, getting angry, hating, breaking, smashing, burning, destroying, scorning, expressing contempt, expressing disdain, cursing, swearing, threatening, insulting, belittling, repudiating, ridiculing.

Fantasy aggression scores ranged from 1 to 15, with a mean of 5.3. The corrected matched-half reliability coefficient was .86; the inter-judge scoring reliability coefficient was .92.

Overt Aggression

To measure overt aggression in the child, a modified sociometric device, the "Guess who" technique (7), was adopted. The *S*s were presented with a booklet containing a series of written descriptions of children, and asked to identify each of these descriptive characterizations by naming one or more classmates. Fifteen overt aggression items were used, such as "Here is someone who is always looking for a fight." A diversity of aggressive behaviors were included; items depicted verbal, unprovoked physical, provoked physical, outburst, and indirect forms of aggressive behavior.

An overt aggression score was obtained for each subject by counting the number of times he was named by his classmates. There were substantial differences among the three classes in the distributions of the overt aggression scores; in order to combine into one distribution the scores of children in different classes, overt aggression raw scores were transformed into standard scores.

The biserial correlation coefficient between the overt aggression measure derived from the children

and teacher entries for the same "Guess who" aggression items was .76 ($p < .01$).

RESULTS

Two Pearson product-moment correlation coefficients were obtained. For boys ($N = 23$) whose mothers are relatively encouraging or supportive of aggression, the correlation between fantasy aggression and overt aggression is +.43 ($p < .05$, two-tailed test). For boys ($N = 21$) whose mothers are relatively discouraging of aggression, the corresponding correlation is −.41 ($p < .10$, two-tailed test). These coefficients are statistically different ($p = .006$, two-tailed test).

When the total sample is not separated into two groups on the basis of scores for maternal response to aggression, the overall Pearson product-moment correlation coefficient is +.07. This coefficient is not significantly different from zero.

DISCUSSION

Confirmation is found for the hypothesis that under conditions of relative maternal encouragement of aggression, a greater degree of correspondence exists between the fantasy and overt aggression of children than under conditions of relative maternal discouragement of aggression. Thus, the direction and extent of the relationship between fantasy and overt aggression in the child is apparently influenced by the maternal attitudes and practices surrounding the learning of aggressive behavior.

It has been predicted (13, 16) that those tendencies which are negatively sanctioned or prohibited will be high in fantasy expression and low in overt expression. This association is premised upon a compensatory or substitutive role of fantasy where overt expression is not allowed. A scatter plot of the fantasy and overt aggression scores for the children whose mothers discourage aggression (from which the −.41 coefficient is derived) reveals a considerable number of such high fantasy aggression, low overt aggression scores. However, children with low fantasy aggression and high overt aggression scores are as well represented in this scatter plot as those with high fantasy aggression, low overt aggression scores. Although mothers of children in this group were classified (relative to the others) as discouraging aggression, perhaps certain of them do so ineffectively, and thus allow the child sufficient release of aggressive feelings in overt behavior so that he may not need to express aggression in fantasy. An alternative speculation regarding the concurrence of low fantasy aggression and high overt aggression in the group exposed to maternal discouragement of aggression suggests that a child with strong aggressive needs whose mother prohibits aggression may assign this prohibitory attitude to the adult experimenter and suppress fantasy aggression expressions in the testing situation; yet this child may find avenues for overt expression of aggression among his peers.

In the present study, only one condition related to the learning of aggressive responses and controls was assessed, maternal attitudes and practices. Other possibly critical determinants that remain to be explored include fathers' behavior and teachers' attitudes and practices. This study has sampled a limited range of maternal attitudes and practices concerning aggression. Although there is no direct manner of determining the absolute degree of punitiveness of the most prohibitive mother in this sample, it appears unlikely that extremely severe and continuous maternal punitiveness is represented. Such severe condemnation of aggression might so limit or restrict both the fantasy aggression and overt aggression expressions of the child that no correlational analysis within such a group would be possible. Both the extremes of unimpeded permissiveness and severe condemnation warrant further investigation.

SUMMARY

The relationship between fantasy and overt expressions of aggression was studied as a function of the maternal attitudes and practices toward aggression. Subjects were 44 boys and their mothers. The boys' fantasy aggression was assessed through a modified TAT approach, their overt aggression was measured through a modified sociometric technique, and maternal attitudes and practices toward aggression were measured by use of a questionnaire-interview device.

Support was found for the hypothesis that under conditions of maternal encouragement of aggression, a greater degree of correspondence exists between fantasy and overt aggres-

sion of children than under conditions of maternal discouragement of aggression.

REFERENCES

1. BACH, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, **59**, No. 2 (Whole No. 272).
2. BIALICK, I. The relationship between reactions to authority figures on the TAT and overt behavior in an authority situation by hospital patients. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1951.
3. CHILD, I. L., FRANK, KITTY F., & STORM, T. Self-ratings and TAT: Their relations to each other and to childhood background. *J. Pers.*, 1956, **25**, 98–114.
4. DAVIDS, A., HENRY, A. F., MCARTHUR, C. C., & MCNAMARA, L. F. Projection, self evaluation, and clinical evaluation of aggression. *J. consult. Psychol.*, 1955, **19**, 437–440.
5. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, **50**, 3–11.
6. GLUCK, M. R. The relationship between hostility in the TAT and behavioral hostility. *J. proj. Tech.*, 1955, **19**, 21–26.
7. HARTSHORNE, H., & MAY, M. A. *Studies in the nature of character.* II. *Studies in service and self-control.* New York: MacMillan, 1929.
8. KAGAN, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, **52**, 390–393.
9. KAGAN, J., & MUSSEN, P. H. Dependency themes on the TAT and group conformity. *J. consult. Psychol.*, 1956, **20**, 29–33.
10. KORNER, ANNELIESE F. *Some aspects of hostility in young children.* New York: Grune & Stratton, 1949.
11. LESSER, G. S. Maternal attitudes and practices and the aggressive behavior of children. Unpublished doctor's dissertation, Yale Univer., 1952.
12. LINDZEY, G. Thematic Apperception Test: Interpretive assumptions and related empirical evidence. *Psychol. Bull.*, 1952, **49**, 1–25.
13. MURRAY, H. A. *Thematic Apperception Test Manual.* Cambridge: Harvard Univer. Press, 1943.
14. MUSSEN, P. H., & NAYLOR, H. K. The relationships between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, **49**, 235–240.
15. PITTLUCK, PATRICIA. The relation between aggressive fantasy and overt behavior. Unpublished doctor's dissertation, Yale Univer., 1950.
16. SANFORD, R. N., ADKINS, MARGARET M., MILLER, R. B., COBB, E. A., *et al.* Physique, personality, and scholarship: A cooperative study of school children. *Monogr. Soc. Res. Child Developm.*, 1943, **8**, No. 1.
17. SEARS, R. R. Relation of fantasy aggression to interpersonal aggression. *Child Developm.*, 1950, **21**, 5–6.
18. SYMONDS, P. M. *Adolescent fantasy: An investigation of the picture story method of personality study.* New York: Columbia Univer. Press, 1949.
19. TOMKINS, S. S. *The Thematic Apperception Test.* New York: Grune & Stratton, 1947.

Received November 2, 1956.

AFFILIATION MOTIVATION AND PRODUCTIVITY IN SMALL GROUPS

RICHARD DECHARMS[1]

University of North Carolina

THE general hypothesis underlying the research reported here is that there is a systematic relationship between the strength of affiliation motivation and cooperative group behavior. At first glance one might expect that affiliation motivation would facilitate cooperative group behavior and inhibit competitive behavior. It seems plausible that a person high in the affiliation motive should perceive working for a group goal as an activity instrumental to acceptance by the group, and therefore be predisposed to cooperate rather than compete.

These simple predictions are complicated by at least two factors. First, it has been shown (1, 5) that the experimental arousal of motives depends on the conditions under which *S* works. It follows that the predictions should hold only under certain arousal conditions. A second factor has been pointed out by French and Chadwick (2) who found that affiliation motivation can be predominantly goal-oriented or predominantly threat-oriented, and that these different orientations may have different behavioral concomitants. For instance, unpopular *S*s tend to show more threat-oriented responses on a projective measure of motivation.

Threat-oriented and goal-oriented motivation probably result from differential experiences of acceptance in the past. The goal-oriented person redintegrates positive past experiences of acceptance in an affiliation situation, while the threat-oriented person redintegrates negative experiences of rejection. The person who has been accepted in the past may have learned that cooperation leads to further acceptance. The previously rejected person, on the other hand, may desire acceptance so strongly that anticipation of rejection is particularly frustrating and leads to an aggressive competitive response. Affiliation motivation of the threat-oriented type may therefore facilitate competition rather than cooperation. Competition may in turn have been rewarded by recognition, which may even have led to greater acceptance of the originally rejected person.

These considerations lead to the following hypotheses:

Hypothesis I. Under conditions that lead them to anticipate rejection, *S*s high in threat-oriented affiliation motivation are more productive on a competitive task than *S*s low in threat-oriented affiliation motivation.

Hypothesis II. Under conditions that lead them to anticipate rejection, *S*s high in threat-oriented affiliation motivation are less productive on a cooperative task than *S*s low in threat-oriented affiliation motivation.

Under conditions that lead them to anticipate acceptance, *S*s high in goal-oriented affiliation motivation should be more productive on a cooperative task and less productive on a competitive task than *S*s low in goal-oriented affiliation motivation. The present experiment cannot adequately test these predictions concerning goal-oriented affiliation motivation, since the design did not include situations in which anticipation of acceptance was made salient. However, the effects of goal-oriented affiliation motivation may appear in a weaker form in the control groups which experienced no arousal of anticipated acceptance or rejection.

METHOD

Subjects

A total of 88 male undergraduates participated in the experiment. They were selected from several sections of an introductory psychology course according to their affiliation and achievement motivation scores, (n Affiliation, n Achievement) and met in groups of four, five, or six for the experiment. Individual groups were composed of *S*s both high and low in n Affiliation.

[1] This report derives from a larger study submitted to the University of North Carolina as a doctoral dissertation. The author is indebted to Drs. John W. Thibaut, Edward E. Jones, and to the other members of the Organization Research Group for constructive criticism.

The study was conducted in the program of the Organization Research Group of the Institute for Research in Social Science, which is supported jointly by the Ford Foundation and the Office of Naval Research [Nonr-855 (04)].

The selection procedure assured comparable n Achievement scores for *S*s in the various cells of the design.

Design

In each experimental condition, the actual tasks to be performed by *S*s were identical. Only the task instructions and the experimental procedures prior to the task were manipulated. Half of the groups received competitive task instructions; the other half, cooperative task instructions. The experimenter (*E*) attempted to arouse the affiliation motive and anticipation of rejection in half of the groups; for control, this procedure was omitted with the other half. Thus, the experimental procedures called for a two-by-two design, varying task instructions and arousal conditions. Within each of the four cells, *S*s could be divided into those above and those below the median of the measures of motivation.

Measurement of n Affiliation

During their first laboratory meeting the *S*s were asked to write stories in response to picture stimuli according to the procedure described by McClelland *et al.* (5). A six-picture measure of n Affiliation was obtained by scoring according to the procedure described by Atkinson *et al.* (1). Five of the pictures have been used by McClelland and his associates. The sixth picture, representing four men presumably discussing something or chatting casually, was taken from Henry and Guetzkow (3).

Each story containing affiliation imagery was classified as either predominantly positive (goal-oriented) or predominantly negative (threat-oriented). A story was classified as positive when it contained predominantly approach responses, such as statements about positive affective relationships between characters, or characters striving for and attaining affiliation goals (positive outcome). Components considered positive were companionate activities, such as parties, reunions, visits, small talk, bull sessions, etc. A story was classified as negative when it contained predominantly avoidance responses, such as statements about concern over separation, or characters striving for an affiliation goal not specifically attained in the story (negative or questionable outcome). Negative components included concern over: being left out or rejected; being lonely or without friends; experiencing physical separation including death; experiencing psychic separation, such as disagreements or quarrels (including attempts to make amends for past actions that brought about separation or rejection); and experiencing no reciprocation of love.

The +Affiliation score of any *S* was the total of the n Affiliation scores for all stories which were judged to be predominantly positive. The −Affiliation score was derived in a corresponding manner. The sum of the positive and negative affiliation scores for an *S* (without regard to sign) thus equals the total n Affiliation score as described by Atkinson *et al.* (1).

The pictures were found to elicit approximately equal numbers of positive and negative stories and thus appear to yield a balanced measure of total n Affiliation. There was essentially no correlation between the positive and the negative affiliation motivation scores ($r = -.15$, $N = 88$).

Interscorer reliability of two experienced scorers[2] was measured throughout the scoring by spot checking approximately one-third of all stories scored. The agreement between scorers was 95 per cent for affiliation imagery alone, and 92 per cent for all categories.

Affiliation Arousal

Each *S* who underwent this procedure was asked to rank a series of traits individually from most to least complimentary. An eight-minute group discussion followed in which the group was to try to agree on a single group ranking of the traits. Each *S* was told that he was also to form an opinion of the other members, since he would be asked to evaluate them and to choose work partners for a subsequent task. After the discussion, each *S* stood before the group and read his original ranking of the traits, while all others chose two characteristic traits for him. Finally, each *S* indicated his preference for work partners by ranking the others. This entire procedure was omitted in the control groups.

The arousal procedure followed closely that described by Atkinson *et al.* (1) with the addition of the group discussion. This was included in order to give the *S*s some basis for rating each other and so that some of the characteristics of a group could begin to emerge. The ratings and partner choices were intended to make *S* acutely aware of being evaluated by other members of the group. Both Shipley and Veroff (6) and Atkinson *et al.* (1) point out that this brings out fear of rejection. Therefore, *S*s in the affiliation arousal condition may be considered to anticipate rejection much more strongly than those in the control condition.

Task Instructions

The competitive instructions presented an individual goal of competition for highest productivity, the results to be a ranking of individuals in the group. The cooperative instructions presented a team goal of cooperation among members on individual problems to be summated for the group score.

To make the cooperative task instructions more realistic, the group was instructed to divide the total task among themselves in a specified way so as to finish as many of the problems as possible, each person starting on a different page in the booklet, working the remaining pages in order, and returning to the first page if he reached the end. It was stressed that he should do as many problems as possible since this would increase the group score. The booklets were actually composed so that although each *S* started on a different page, they were all actually working on the same problems in the same order. The task was therefore comparable under all conditions.

Group Attractiveness

On the post-experimental questionnaire, each *S* was asked to rate on a seven-point scale his agreement with items concerning his liking for the group during the experiment. The items were as follows: (*a*) If I had to choose a group to work with me on a real problem, I would choose some of the members of this group. (*b*) I

[2] Mrs. John S. Crittenden and the author.

TABLE 1

MEAN NUMBER OF TASKS ATTEMPTED AND THREAT-ORIENTED AFFILIATION MOTIVATION

Task and Level of Threat-Oriented Affiliation Motivation	Mean Tasks Attempted	
	Affiliation Arousal (Threat of Rejection)	No Arousal
Competitive Task		
High −Affiliation	133.8	126.3
(*n*)	(10)	(8)
Low −Affiliation	112.5	115.0
(*n*)	(12)	(14)
Mean	122.2	119.1
(*n*)	(22)	(22)
Cooperative Task		
High −Affiliation	100.6	87.8
(*n*)	(11)	(10)
Low −Affiliation	110.2	84.6
(*n*)	(11)	(10)
Mean	105.4	86.3
(*n*)	(22)	(22)

TABLE 2

MEAN NUMBER OF TASKS ATTEMPTED AND GOAL-ORIENTED AFFILIATION MOTIVATION

Task and Level of Goal-Oriented Affiliation Motivation	Mean Tasks Attempted	
	Affiliation Arousal (Threat of Rejection)	No Arousal
Competitive Task		
High +Affiliation	118.4	118.3
(*n*)	(9)	(14)
Low +Affiliation	124.8	120.5
(*n*)	(13)	(8)
Cooperative Task		
High +Affiliation	101.8	92.5
(*n*)	(12)	(13)
Low +Affiliation	109.7	77.4
(*n*)	(10)	(9)

felt some dislike for at least one of the members during the tasks. (Reverse scored.) (*c*) In general, I liked most of the people in my group very well.

Productivity

The productivity score for each *S* consisted of the sum of the problems attempted on both tasks. Task I was a simple arithmetic task. Task II was a scrambled words task. Performance on both of these tasks has previously been found to be related to achievement motivation (4, 7).

Procedure

The affiliation arousal was the first procedure for aroused groups and was omitted for control groups. The task instructions placing *S* in a competitive or a cooperative task situation were then read, after which *Ss*, seated in isolation booths, worked on 120 arithmetic problems for 12 minutes, followed by 12 minutes' work on 120 scrambled words. Feedback was supplied each *S* privately, informing him that his productivity ranking was just below the median rank of his group. This was done immediately prior to administration of the post-experimental questionnaire in order to equate effects of knowledge of results on the tasks. The questionnaire included the "Group Attractiveness Scale" and concluded the experiment except for an informal discussion with the *Ss*.

RESULTS

Table 1 presents the productivity means for *Ss* high and low in −Affiliation under the four different experimental conditions. Considering only the experimental manipulations, analysis of variance demonstrates that mean productivity in the competitive task situation was significantly higher than mean productivity in the cooperative task situation ($p < .01$). There was no significant difference between mean productivity under affiliation arousal and under no arousal. However, considering only the cooperative task groups, mean productivity of *Ss* under affiliation arousal was significantly higher than productivity under no arousal ($p < .01$).

A median split of +Affiliation or −Affiliation scores within each of the four experimental conditions resulted in unequal cell entries. A multiple regression analysis was therefore computed (see Table 3) to determine the effects due to level of motivation within each experimental condition. For descriptive purposes the means in Tables 1 and 2 for high and low levels of motivation will be discussed. However, significance levels will refer to tests of the variance due to regression.

Hypothesis I predicts higher productivity for *Ss* high in −Affiliation than for *Ss* low in −Affiliation on a competitive task under affiliation arousal (threat of rejection). Table 1 shows a large difference in the predicted direction, and regression analysis reveals that this is a reliable difference ($p < .01$) confirming the hypothesis. Under no arousal there was a similar but insignificant difference.

The data in Table 1 are also consistent with Hypothesis II which predicts lower productivity for *Ss* high in −Affiliation than for *Ss* low in −Affiliation on a cooperative task under affiliation arousal. The regression analysis, however, reveals that this difference falls short of customary standards of significance ($p < .07$).

TABLE 3
ANALYSIS OF VARIANCE DUE TO REGRESSION OF PRODUCTIVITY ON GOAL-ORIENTED AND THREAT-ORIENTED AFFILIATION MOTIVATION SCORES

Arousal	Task	Motive	*df*	Mean Square	*F*	*b*
Affiliation Arousal	Competitive	+Affiliation	1	814	1.15	−2.1754
		−Affiliation	1	5,173	7.30***	+6.0783
	Cooperative	+Affiliation	1	547	.77	−1.7836
		−Affiliation	1	2,546	3.59**	−4.2643
No Arousal	Competitive	+Affiliation	1	110	.16	−0.8012
		−Affiliation	1	490	.69	+1.8714
	Cooperative	+Affiliation	1	2,027	2.86*	+3.4327
		−Affiliation	1	101	.14	+0.8500
	Error†		76	709		

* $p < .10$.
** $p < .07$.
*** $p < .01$.
† Twelve degrees of freedom were used in testing variance due to differences between cell means and variance due to regression.

For a test of the significance of the interaction of level of threat-oriented n Affiliation with kind of task under affiliation arousal, the difference between the two computed regression coefficients was tested. The significant relationship is found that would be predicted by combining Hypotheses I and II ($p < .01$).

It has been noted that the experimental design did not provide a situation of heightened anticipation of acceptance for testing predictions for +Affiliation, but data for the No Arousal condition provide suggestive evidence. Table 2 presents means for *Ss* high and low in +Affiliation. As predicted, productivity on a cooperative task under No Arousal was higher for *Ss* high in +Affiliation than for *Ss* low in +Affiliation, although the trend falls short of statistical significance ($p < .10$).

Analysis of productivity and total n Affiliation scores reveals smaller relationships similar to those with −Affiliation in the arousal cells, but +Affiliation and −Affiliation scores considered separately seem to be better predictors of performance than total n Affiliation under the present conditions.

DISCUSSION

The main hypotheses, which are in general supported by the data in Table 1, were based on the assumption that competitive productivity under certain conditions is the result of hostility induced in *Ss* high in −Affiliation by threat of rejection. Using the Group Attractiveness Scale as a rough measure of hostility toward the group, it was possible to examine the relationships between hostility, −Affiliation, and the experimental manipulations. Although the relationships were all small and insignificant, they were remarkably consistent. Hostility was greater in Arousal than in No Arousal cells, and greater for *Ss* high in −Affiliation than for *Ss* low in −Affiliation. Competitive productivity followed the same pattern. It was higher in the Arousal cells, and (as shown in the test of Hypothesis I) high for *Ss* high in −Affiliation. Further, competitive productivity was higher for *Ss* high in hostility than for *Ss* low in hostility; and, conversely, cooperative productivity was lower for *Ss* high in hostility than for *Ss* low in hostility.

Table 1 also shows that cooperative productivity was significantly greater under Arousal than under No Arousal. This finding cannot be attributed to threat of rejection, which should inhibit cooperative productivity, especially in *Ss* most influenced by threat. It may well be attributable to the group discussion, in which only the aroused group participated and which allowed *Ss* to become acquainted. They seemed to gain a more positive attitude toward the situation, as indicated by post-experimental discussions with the groups, and this positive attitude could well have produced an increased incentive for cooperative work.

Further examination of the data shows that cooperative productivity under arousal was relatively greater for *Ss* low on any of the measures of affiliation motivation. The most striking difference was with total n Affiliation. Cooperative productivity for *Ss* low in n

Affiliation was higher in the aroused than in the control condition, (aroused mean = 116.8, control mean = 79.5), while productivity for *Ss* with high n Affiliation was approximately the same under the two conditions (aroused mean = 89.0, control mean = 90.5). The overall increase under arousal, which is attributable entirely to the *Ss* with low motivation, may be similar to Wendt's findings (7) with achievement motivation under scheduled and nonscheduled task conditions. He interprets his results as showing the effect of imposing extrinsic motivation upon *Ss* low in intrinsic motivation in the scheduled situation, thereby bringing their productivity up to those who were intrinsically motivated and had high productivity under either scheduled or nonscheduled conditions. In the present experiment, however, the productivity under arousal of *Ss* low in n Affiliation was significantly higher than that of the *Ss* high in n Affiliation. It would seem that the extrinsic threat of rejection has interacted with the intrinsic n Affiliation, thereby inhibiting increased cooperative productivity under the arousal condition for *Ss* high in n Affiliation, while the group discussion facilitated cooperative productivity for *Ss* low in n Affiliation to whom the threat was not as important.

SUMMARY

This study tested the hypotheses that under conditions favoring the anticipation of rejection, *Ss* high in threat-oriented affiliation motivation are (*a*) more productive on a competitive task, and (*b*) less productive on a cooperative task than *Ss* low in threat-oriented affiliation motivation. These hypotheses were in general supported by data from 88 undergraduate *Ss*.

In the control condition, which was treated neutrally with respect to arousal of anticipated acceptance or rejection, *Ss* high in goal-oriented affiliation motivation (as compared to *Ss* low in this respect) tended to be more productive on a cooperative task. This relationship did not reach statistical significance.

Under Affiliation Arousal (which included both a group discussion and threat of rejection) there was higher cooperative productivity than under No Arousal, a difference contributed almost entirely by *Ss* low in total affiliation motivation (goal-oriented plus threat-oriented). This result was interpreted as showing an interaction between the effects of the group discussion and threat of rejection. The group discussion facilitated cooperative productivity in general, but among *Ss* high in affiliation motivation this effect was counteracted by the threat of rejection.

The other relationships between overall affiliation motivation and productivity were insignificant. The importance of distinguishing between goal- and threat-oriented affiliation motivation was stressed.

REFERENCES

1. ATKINSON, J. W., HEYNES, R. W., & VEROFF, J. The effect of experimental arousal of the affiliation motive on thematic apperception. *J. abnorm. soc. Psychol.*, 1954, **49,** 405–410.
2. FRENCH, ELIZABETH, G., & CHADWICK, IRENE. Some characteristics of affiliation motivation. *J. abnorm. soc. Psychol.*, 1956, **52,** 296–300.
3. HENRY, W. E., & GUETZKOW, H. Group projection sketches for the study of small groups. *J. soc. Psychol.*, 1951, **33,** 77–102.
4. LOWELL, E. L. The effect of need for achievement on learning and speed of performance. *J. Psychol.*, 1952, **33,** 31–40.
5. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive.* New York: Appleton-Century-Crofts, 1953.
6. SHIPLEY, T. E., JR., & VEROFF, J. A projective measure of need for affiliation. *J. exp. Psychol.*, 1952, **43,** 349–356.
7. WENDT, H. W. Motivation, effort, and performance. In D. C. McClelland (Ed.), *Studies in motivation.* New York: Appleton-Century-Crofts, 1955. Pp. 448–459.

Received October 25, 1956.

COGNITIVE SET AND MOTIVATIONAL FACTORS IN THE PERCEPTION OF NEUTRAL AND THREAT-RELATED STIMULI[1]

RUTH C. WYLIE

Sarah Lawrence College

RECENTLY, experimental psychologists have been interested in the proposition that perceptual thresholds are higher for so-called emotionally threatening stimuli, and lower for positively valued stimuli than for affectively neutral ones. Tests of such a proposition have yielded ambiguous results because of difficulties in controlling or knowing the status of such factors as: (*a*) subjects' attitudes toward the stimuli, (*b*) frequency of subjects' past experiences with the stimuli (*c*) subjects' cognitive sets or expectancies; and (*d*) the availability of the required responses as they are determined by subjects' previous practice in making the responses and willingness to report honestly (1–3; 5–11). This experiment employed new means to control the above-mentioned variables in a comparison of the influence of motivational factors plus cognitive set with the influence of cognitive set alone. In this design the experimenter (*E*) induced the desired emotional state and cognitive expectancies in the subjects under laboratory control; and the use of pretested numbers as stimuli, presented under specially devised conditions, provided improved control of factors *b*, *c*, and *d*.

The following hypotheses were tested:

Hypothesis 1. When subjects (*S*s) have just been given an auditory presentation of a stimulus and have reason to expect that they may be given the same stimulus visually, their duration thresholds for such visual stimuli are lower than for otherwise comparable visual stimuli for which no such cognitive set has been established.

Hypothesis 2. When *S*s desire to succeed in a task, their duration thresholds for visual stimuli indicating that they have failed are greater than their duration thresholds for otherwise comparable visual stimuli indicating that they have succeeded. These differences are expected to be greater than could be accounted for on the basis of cognitive set or expectancy alone.

Hypothesis 3. When *S*s desire to succeed in a task, the content of their prerecognition guesses about visual stimuli which indicate that they have failed resembles the content of stimuli which, had they been presented to *S*, would have indicated success. This distortion of prerecognition guesses is expected to be greater than could be accounted for on the basis of cognitive set or expectancy alone.

METHOD

Stimulus Materials

Two-place numbers were used as stimuli on the assumptions that: (*a*) such stimuli would be emotionally neutral to *S*s until experimental conditions produced an emotional set toward them; (*b*) the response availability for all two-place numbers would be the same, insofar as this is determined by frequency of past visual exposure, practice in pronunciation, and lack of embarrassment in pronunciation. On the basis of a preliminary experiment, testing the tachistoscopic recognizability of two-place numbers of a standard print type and size, these numbers were classified into three groups: hard, medium, and easy to recognize, for use in matching the "threatening" and "rewarding" numbers in the main experiment.

Subjects

The *S*s were 51 volunteer female college students, naive with respect to the purpose of the experiment. They were divided into three groups of 17 each.

Procedure

Group I. Group I *S*s were told they were to participate in two separate experiments. In the "first experiment" (Task 1), each *S* was given individually a specially devised and pretested task. She was told that, although the parts of the task were simple, this was a new intelligence test which revealed the ability of a person to keep her head and concentrate in a fast-moving, distracting situation which might "rattle" her. She was instructed that preliminary information indicated that this is a very important kind of intelligence for everyday living, and she was encouraged to keep trying by being told she would receive partial credit for nearly correct answers. This task consisted of 14 chains of arithmetic operations, with ten operations per chain. The *E* read each chain aloud at the rate of about one operation every five seconds, while *S* did the

[1] The Gerbrands tachistoscope used in this study was purchased under a grant from the American Philosophical Society. Thanks are due also to Dr. Georgiana Wylie who pretested the ego-involving task at another college.

operations mentally and wrote only her final answer for each chain. The oral presentation and mental computation prevented *S* from knowing whether any answer was right or wrong unless she had completely failed to give an answer or had obtained a number with more than two places. The *S* was asked to guess how many problems she had done correctly and how many almost correctly; and to tell how sure she was of this guess. (No *S* was sure of her guess.) Each *S* was asked to rate herself on her ability on this type of task, as compared with the average college student. Finally, she was given fictitious norms which were higher than the actual attainment of the college group. Interviews at the end of the entire experimental procedure indicated that *S*s were ego-involved on Task 1.

After Task 1, *S* was told that the "first experiment" was finished and she was given a brief rest period before the "second experiment." Meanwhile *E* prepared, in another room, a series of 14 numbers, to be presented in a Gerbrands mirror tachistoscope. Regardless of *S*'s answers on Task 1, half the numbers for this series were selected to agree and half to disagree with *S*'s answers. The disagreeing numbers were chosen to be at least as readable as the agreeing ones. Agreeing and disagreeing items were put in random sequence to control for practice and fatigue.

The *S* was then brought into the tachistoscope room and asked if, while she was in the laboratory anyway, she would mind helping *E* test a new apparatus. An effort was made to avoid ego-involving *S* in being very accurate in the perceptual task as such, and to keep her from suspecting the connection between Tasks 1 and 2 by telling her that *E* was only testing the lighting conditions and apparatus for use in a further experiment, and was merely trying to make this perceptual task less boring to *S* by using for stimuli the correct answers to Task 1. (The purpose of avoiding ego-involvement in perceptual accuracy was to prevent masking possible motivationally-induced distortions, something which might happen if the accuracy set for the perceptual task were too high.) As part of the detailed instructions, *S* was told to report what she saw on the screen after each successive exposure, basing her guesses on the visual characteristics of the stimuli, *not* on any other method of guessing. The *S* was given a 45-sec. practice period in reading aloud the numbers 0 through 9, in various scrambled orders. She never removed her eyes from the eyepiece during the instructions and practice period, thus achieving adaptation to the level of illumination used during the experiment.

Before the exposure of the first stimulus number from the experimental list, *E* read aloud *S*'s answer to Problem 1 on Task 1 and told her that if the first tachistoscope number was the same as her answer, this would mean that her answer to Problem 1 was correct. Beginning at .01 second, *S* was given successively longer exposures of the number until she made two successive veridical reports, the first of which was taken as the threshold. The duration thresholds in hundredths of a second and all prerecognition guesses were recorded by *E*. This procedure was repeated for each of the remaining 13 numbers.

At the end of Task 2, the entire experiment was explained to *S* and she was interviewed to see if she had become ego-involved and to find out whether she had discerned the true connection between Tasks 1 and 2. The *S*s reported that the directions had been convincingly misleading, but some of them said they had wondered whether hearing the number and/or wishing to see it might have influenced their perceptions. However, they all reported trying hard to base their guesses exclusively on the visual appearance of the stimuli. Each *S* was asked to keep the experiment secret so that future *S*s would be naive.

It can be seen from the description of the procedure used with Group I that both expectancy and "wishful thinking" might influence their perceptual thresholds for those tachistoscope numbers which agreed with the numbers *E* read out to them.

Group II. Each Group II *S* was arbitrarily paired with a Group I *S*. The treatment given to Group II *S*s was intended to make their experiences in the tachistoscope room exactly the same as those of Group I, except that their attitudes toward the numbers used as visual stimuli should be emotionally neutral, allowing one to study the influence of cognitive expectancies in the absence of motivational influences. Consequently, these *S*s heard nothing about Task 1, but went directly to the tachistoscope room where they were given the same instructions for Task 2 as had been given to Group I, with the following exception: Group II *S*s were told that *E* would "read out a number" instead of that *E* would "read out your answer to a problem on the intelligence test you just took." Any given Group II *S* was given the same list of "read out" and "viewed" numbers as had been used for her corresponding Group I *S*. Each Group II *S* was told, "I can't tell you how many of the numbers flashed on the screen will be the same as those on my list and how many will be different. There might be any degree of agreement between zero and 100 per cent. The greatest likelihood is that there will be about ____ agreements out of 14 numbers which you will have flashed on the screen. You can't be sure of this, however." The number of agreements suggested to a given Group II *S*, was the number of answers the corresponding Group I *S* stated she thought she had gotten correct at the end of her Task 1 test.

A careful inspection of the procedures given Groups I and II suggested that any differences in performance between these groups were to be interpreted with caution, because the parallel instructions to the groups did not suffice completely to equate cognitive expectancies between them. The *S*s in Group I, for example, had reason from Task 1 to *expect* (as well as to hope) that some of the numbers might be *nearly* agreeing, if not exactly agreeing, even though the tachistoscope directions did not mention this possibility. On the other hand, the Group II *S*s had had no comparable previous experience that might have established a cognitive expectancy of a certain number of *near* agreements. In an attempt to equate this factor of cognitive expectancy more effectively, the procedure used for Group II was administered to a third group, with the single modification that the directions were changed by the addition of the underlined clause below: ". . . . The greatest likelihood is that there will be about____ agreements out of 14 numbers which you will have flashed on the screen, *with perhaps____ coming fairly close*." The exact number mentioned in the underlined clause was that which the corresponding Group I *S*

thought she had gotten nearly correct in Task 1. By this change of directions the expectancy of number of *near* as well as exact agreements was made more comparable between each Group III *S* and her matched Group I counterpart.

RESULTS

For each *S*, the median threshold for all agreeing numbers was found (called *A*) and the median threshold for disagreeing numbers was obtained (called *D*).[2] (*A* − *D*) was predicted to be negative in the cases of all three groups. Groups II and III were employed to test Hypothesis 1 which was confirmed by the finding that 31 out of 34 *S*s in these groups showed negative values for (*A* − *D*), a proportion significant far beyond the .01 level. Since other variables have been controlled, this result can reasonably be attributed to the effects of hearing the number and expecting the possibility of seeing it, i.e., to cognitive sets.

Although simply hearing the number and expecting possible agreement leads to faster perception, the second hypothesis asserts that *hoping* to see what one has just heard and *hoping not* to see any other number should make the (*A* − *D*) values of the Group I *S*s more negative than those of Group II or Group III *S*s. In order to make such comparisons, a refined measure of the discrepancy between thresholds for agreeing and disagreeing items was needed that takes into account the individual differences in intra-individual variability in thresholds, in terms of which the groups were not equated. To obtain such a measure, an index of intra-individual variability (here called *I*) was first computed for each *S* by taking the range of her middle six thresholds.[3] Her (*A* − *D*) value could then be evaluated in relation to her overall intra-individual variability. The score used as the dependent variable in testing Hypothesis 2 may be summarized thus:

$$\frac{(A - D)}{I}$$

where

A = *S*'s median threshold for agreeing numbers
D = *S*'s median threshold for disagreeing numbers
I = range of middle six thresholds for *S*

The difference in mean score between Group I and Group II was found to be +0.11, which yielded an insignificant *t* value of 0.60[4], while the difference in mean score between Group III and Group I was only −0.01. Thus, the data give no support to Hypothesis 2; i.e., there is no indication that wishful thinking added to expectancy produces any greater effect on duration thresholds than does expectancy alone.

The third hypothesis postulated that prerecognition guesses of the Group I *S*s might reveal "wishful thinking" by tending to be closer to the "read-out" number than were the prerecognition guesses of Group II or Group III *S*s. To test this hypothesis, all prerecognition guesses within ±10 of the "read-out" number (*S*'s answer in the case of Group I), were tallied and compared to the total number of prerecognition guesses made by *S*, yielding a proportion score for each *S*. The differences between the proportion score of each Group I *S* and that of her corresponding Group II *S* were obtained by subtraction across pairs. The mean difference was 0.14 and $t = 2.48$, which is significant between the .02 and .01 level on a one-tailed test. These results support the hypothesis that the Group I *S*s, who are hoping to see a number at least *near* their answer, will tend to see more numbers near the "read-out" number than will the control *S*s, who have no reason to care whether the tachistoscope number is near the "read-out" number. However, for reasons described under Method, a comparison of Group III with Group I provides a better controlled test of Hypothesis 3. When the prerecognition guesses of Group III were compared to those of Group I in the same way

[2] An item on which a Group I *S* knew she had failed because she got no answer or got more than a two-place number for an answer was eliminated from the calculation of *D* of both the Group I *S* and the corresponding Group II *S*. To keep the counterbalancing of the ordinal position of agreeing and disagreeing items intact, in such a case, the nearest agreeing item was also eliminated from the calculations of *A* for the Group I and corresponding Group II *S*.

[3] This measure of variability was used rather than the *SD* because: (*a*) the number of thresholds in each individual's distribution was small, (*b*) the individual distributions were often skewed. (*I* approximates the interquartile range without involving interpolations or decimal values unwarranted by the size of the distributions or precision of the data.)

[4] No correlation on the dependent variable was found between Group I and Group III *S*s, so uncorrelated *t* formulae were used.

that II and I had been compared, the difference was +.08, which favored the hypothesis at the .10 level (one-tailed test). Groups II and III did not differ significantly between themselves as to prerecognition guesses, so they were combined and then compared with Group I. The difference was +.11, with a P slightly $> .05$ on a one-tailed test.

DISCUSSION

The results of this experiment give clear-cut evidence for the effects of cognitive expectancies upon duration thresholds, whether or not motivational influences which might operate in the same direction were present. The data give some support to the notion that prerecognition guesses may be modified by motivational influences such as desire to succeed in an ego-involving task. No support is given the proposition that such motivational influences affect duration thresholds when other variables are held constant (the usual definition of perceptual defense).

The question arises as to whether those differences which were found are attributable to effects on the perceptual processes or to effects on the response processes, to follow the distinction developed most recently by Garner, Hake, and Ericksen (4). It seems likely that the effects are on the perceptual processes in view of the operations instituted to control for the influences on response processes. In the first place, the use of numbers should make motivational reasons for withholding responses less than they would be in the case of taboo words. In addition, the choice of two-place numbers provided stimuli which the *S*s had had much practice in pronouncing. And finally, the instructions to base guesses solely on the visual properties of the flash should tend to reduce guessing behavior from other sources.

The results on cognitive set are in line with much previous work, and they further point up the need for controlling for this factor carefully in all studies where it is not the independent variable.

The data on the role of motivation are of interest in the light of the current controversy concerning the status of perceptual defense and accentuation. A number of investigators have suggested that the positive experimental findings in this area are attributable to uncontrolled variables such as set, frequency of *S*'s past experience with the stimuli and the like. The effects obtained are sometimes diminished or even reversed when these other variables are altered (6, 10, 11). In one well controlled recent experiment, no evidence of perceptual defense could be obtained (13). It was hoped that the new methods and stimulus materials introduced in this study would also effectively control the variables which have led to ambiguity in previous studies. Insofar as this aim has been realized, these results give no evidence of motivational effects on duration thresholds under comparable conditions of cognitive expectancy.

The present findings, and those of other experiments, may perhaps be reconciled with the concept of perceptual defense on the assumption that perceptual defense occurs only within certain limiting conditions. So far, these conditions remain to be explored and specified precisely, and one value of this experiment lies in suggesting what some of them may be. First, it is possible that laboratory-manipulated ego-involvement, while desirable for precision of control, may produce motivation too weak to affect duration thresholds. Since the motivation was apparently strong enough to exert some effects on the prerecognition guesses, one might infer that different minimum strengths and/or durations of motivation may be required to affect different measures of perceptual distortion.

It might be argued, second, that motivational effects do not appear when some other, more influential variable such as cognitive set is present to override them. On this view, the present results could be interpreted as analogous to the findings of Postman and Schneider (12) on familiarity, from which they concluded that differences in value rank produced significant differences in perceptual sensitivity to unfamiliar but not to highly familiar words.

Finally, there is the question whether the findings on perceptual defense may vary as a function of types of material used. Perhaps affect can become attached to some stimuli (e.g., words and pictures) more easily than to numbers. If so, the psychological variables leading to such differences among materials need to be determined.

In any event, the present findings suggest that caution and further empirical qualifications are in order before making generalizations about the occurrence of perceptual

defense, and about the variables which limit it or with which it interacts, if and when it occurs.

SUMMARY AND CONCLUSIONS

An experiment testing the following three hypotheses was described:

1. When *Ss* have just been given an auditory presentation of a stimulus and have reason to expect that they may be given the same stimulus visually, their duration thresholds for such visual stimuli are lower than for otherwise comparable visual stimuli for which no such cognitive set has been established.

2. When *Ss* desire to succeed in a task, their duration thresholds for visual stimuli indicating that they have failed are greater than their duration thresholds for otherwise comparable visual stimuli indicating that they have succeeded. These differences are expected to be greater than could be accounted for on the basis of cognitive set or expectancy alone.

3. When *Ss* desire to succeed in a task, the content of their prerecognition guesses about visual stimuli which indicate that they have failed resembles the content of stimuli which, had they been presented to *S*, would have indicated success. This distortion of prerecognition guesses is expected to be greater than could be accounted for on the basis of cognitive set or expectancy alone.

The design attempted to control by experimental manipulation variables which have caused difficulty in previous experiments in this field, i.e., *S*'s attitude toward the stimulus materials, frequency of his past experience with perceiving the stimulus materials and pronouncing them, inherent readability of the stimulus forms and willingness to report honestly.

The influence of cognitive set on duration thresholds was demonstrated (Hypothesis 1). No evidence was found for the effects of motivation on duration thresholds (Hypothesis 2), but some evidence was found for the influence of motivation on the content of prerecognition guesses (Hypothesis 3).

REFERENCES

1. COWEN, E. L., & BEIER, E. G. Threat, word frequencies, and recognition. *J. abnorm. soc. Psychol.*, 1954, **49**, 178–182.
2. ERIKSEN, C. W. The case for perceptual defense. *Psychol. Rev.*, 1954, **61**, 175–182.
3. FREEMAN, J. T. Set or perceptual defense? *J. exp. Psychol.*, 1954, **48**, 283–288.
4. GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. Operationism and the concept of perception. *Psychol. Rev.*, 1956, **63**, 149–159.
5. HOWES, D. On the interpretation of word frequency as a variable affecting speed of recognition. *J. exp. Psychol.*, 1954, **48**, 106–112.
6. LACY, O. W., LEWINGER, NATALIE, & ADAMSON, J. F. Foreknowledge as a factor affecting perceptual defense and alertness. *J. exp. Psychol.*, 1953, **45**, 169–174.
7. LAZARUS, R. Is there a mechanism of perceptual defense? A reply to Postman, Bronson, and Gropper. *J. abnorm. soc. Psychol.*, 1954, **49**, 396–398.
8. NEISSER, U. An experimental distinction between perceptual process and verbal response. *J. exp. Psychol.*, 1954, **47**, 399–402.
9. POSTMAN, L. On the problem of perceptual defense. *Psychol. Rev.*, 1953, **60**, 298–306.
10. POSTMAN, L. The experimental analysis of motivational factors in perception. In M. R. Jones (Ed.), *Current theory and research in motivation: A symposium*. Lincoln, Nebr.: Univer. of Nebraska Press, 1953. Pp. 59–107.
11. POSTMAN, L., BRONSON, W. C., & GROPPER, G. L. Is there a mechanism of perceptual defense? *J. abnorm. soc. Psychol.*, 1953, **48**, 215–224.
12. POSTMAN, L., & SCHNEIDER, B. H. Personal values, visual recognition and recall. *Psychol. Rev.*, 1951, **58**, 271–284.
13. TURKEL, H. Perceptual defense in word recognition: A test of its generalized occurrence. Paper read at Eastern Psychol. Ass. March, 1956.

Received November 12, 1956.

MATERNAL CHILD-REARING PRACTICES AND AUTHORITARIAN IDEOLOGY[1]

I. HART[2]

Duke University

THIS study was an attempt to investigate maternal preferences in techniques for controlling child behavior as a function of socialization situations and a particular aspect of maternal personality. The basic hypothesis was that patterns of child-rearing preferences can be predicted from personality trends expressed in ideological form; specifically, authoritarianism. As a test of this hypothesis, specific predictions were formulated, derived from the clinical insights and descriptive statements of Adorno *et al.* (1) and Levinson and Huffman (4). These predictions were tested in terms of the socialization model of Whiting and Child (7).

Writers on the authoritarian personality have contended that certain child-rearing practices are both antecedents and consequents of authoritarianism. The authoritarian parent is said to emphasize discipline that takes the form of "bodily harm, social isolation and/or shaming, rather than the loss of love, since the giving of love and the establishment of strong affectional ties are not likely to be primary issues in an autocratic setting" (4).

Whiting and Child's analysis of the punishment techniques customarily used by parents in socializing their children bears considerable congruence to the range of behaviors that Levinson and Huffman attribute to high and low authoritarian parents. One phase of Whiting and Child's analysis rests on the assumption that punishments differ in the extent to which they contribute to maintaining the child's orientation to the goal of parental affection. Thus they distinguish three kinds of discipline as "love-oriented," i.e., serving to maintain the child's striving for parental love: denial of love, threats of denial of reward, and threats of ostracism. In contrast to these techniques, they propose three kinds of discipline as having the overall effect of establishing tendencies in the child to avoid the parents. These "nonlove-oriented" punishment techniques are: physical punishment, threats of physical punishment, and punishment by ridicule. Whiting and Child did not include actual denial of rewards and actual ostracism in either the love-oriented or the nonlove-oriented categories because these techniques were considered to contribute potentially both to striving for parental affection and to the development of avoidance tendencies. These latter punishments were described as "ambiguous." On the basis of the descriptions of Adorno *et al.* and of Levinson and Huffman, low and high authoritarian parents might be expected to select love-oriented and nonlove-oriented disciplinary techniques, respectively.

METHOD

One hundred and twenty-six mothers, who had children between the ages of 2½ and 5½, were studied by means of intensive but structured interviews in which each mother was asked to indicate her most probable response to her own child's behavior in specific situations. Free responses were submitted to three independent raters (psychology graduate students) for coding appropriate to the Whiting and Child categories of disciplinary techniques, according to agreement of two out of the three raters. The subjects (*S*s) were white, native born, Christian Americans ranging in age from 23 to 43. Mean mother's age was 30.8; mean age of the children was 4.1.

In an attempt to keep the effects of education and socioeconomic level as constant as possible, the sample was restricted to middle-class mothers who had completed at least one year of college. The fathers were predominantly college graduates, and for the most part were engaged in either professional or higher level managerial occupations. The *S*s were recruited by asking interviewees to suggest mothers who could participate in the study. Almost all of the *S*s were able to suggest others; 95 per cent of those who were asked participated.

The interview schedule contained 38 items depicting various child behavior situations, and a modified version of Whiting and Child's categorization of punishment techniques, combined in questionnaire form. The behavior situations were chosen on the basis of a priori judgment and some brief preliminary research, in the expectation that they would occur rather frequently and tend to elicit some form of discipline rather than

[1] This article is based on a thesis submitted to Duke University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer expresses his appreciation to Dr. E. E. Jones for his helpful criticisms in the design and execution of the study.

[2] Now at Cowell Memorial Hospital, University of California, Berkeley.

reward from parents of children in the age range sampled. It was further intended that each situation represent an instance of behavior in one of the following six areas:

1. Feeding and oral activities: e.g., "insists on eating with fingers," "seldom finishes meal," etc.

2. Cleanliness-toilet training: e.g., "frequently has toilet accidents because won't interrupt play."

3. Sex: sexual curiosity, sexual and obscene remarks, handling genitals, body exposure.

4. Aggression: Items involve insolence and impudence, hitting and pushing, temper tantrums, explicit disobedience, willful property damage.

5. Dependence: These items attempt to depict behavior suggesting that the child lacks self-reliance, is attempting to gain parental support, etc., e.g., "whines or begs for help unnecessarily."

6. Independence: This area includes situations that are intended to suggest that the child is showing initiative, curiosity, assertiveness, e.g., "insists on own way with other children," "wanders away from home," etc.

In addition to the structured interview phase, each mother was asked to describe her three biggest problems concerning her child, to answer two open-end projective-type questions, and to fill out a combined version of the short-form (TFI) scale (4) and F scale (1).[1] The scores thus obtained (M, 3.70; *SD*, 1.09) provided the measure of authoritarian personality. The interviews took from 1½ to 2½ hours.

TABLE 1

MEAN PERCENTAGE OF MOTHERS AT EACH LEVEL OF AUTHORITARIANISM WHO SELECT EACH TYPE OF DISCIPLINE

Type of Discipline and Level of Authoritarianism	Behavior Area						Overall Mean
	Independence	Dependence	Cleanliness	Feeding	Sex	Aggression	
Love							
High	30.33	65.12	59.52	47.29	70.26	18.64	48.53
Medium	27.90	69.52	69.64	58.93	80.57	23.14	54.95
Low	45.05	84.76	88.69	83.31	89.12	42.52	72.24
Mean	34.43	73.13	72.62	63.18	79.98	28.10	
Nonlove							
High	60.55	23.10	33.93	30.57	20.62	69.40	39.69
Medium	53.02	16.45	26.19	20.48	11.17	47.57	29.15
Low	47.60	7.36	8.93	8.81	8.14	41.88	20.45
Mean	53.72	15.64	23.02	19.95	13.31	52.95	
"Ambiguous"							
High	9.12	11.78	6.55	22.14	9.12	11.96	11.78
Medium	19.08	14.03	4.17	20.59	8.26	29.29	15.90
Low	7.35	7.88	2.38	7.88	2.74	15.60	7.31
Mean	11.85	11.23	4.36	16.87	6.71	18.95	

RESULTS

Authoritarianism was found to be unrelated to each of the following variables: age of the mother, age of the child, sex and birth order of the child, and number of children in the family. The behavior situations included in the interview schedule were quite comparable to and representative of the kinds of behavior problems reported by the *S*s.

The main results are described below with reference to three experimental predictions.

1. *The selection of nonlove-oriented disciplinary techniques should be a positive function of maternal authoritarianism score.* To test this prediction, a product-moment correlation was computed between each mother's authoritarianism score and the number of situations for which she selected nonlove-oriented techniques. The correlation thus obtained was .63 ($p < .001$).

For a more detailed analysis, the distribution of authoritarianism scores was divided into thirds, with cutoff scores for highs and lows falling at 4.1 and 3.2, respectively. The relationship between mothers' authoritarianism and preferred disciplinary techniques may be observed in the right-hand marginal column of Table 1. The body of Table 1 lists the percentages of type of discipline selected by mothers of a given authoritarianism level, in the various behavioral areas. Mothers at all levels of authoritarianism in general selected more love-oriented than nonlove-oriented control measures. However, it will also be seen that as authoritarianism level increases, nonlove-oriented techniques are preferred increasingly, and that as authoritarianism level decreases, the tendency to select love-oriented techniques increases. Since practically all of the previous research concerning authoritarianism has employed only the extremes of the distribution, the confirmation of this prediction also lends support to the contention that authoritarianism is an intrinsic continuum, not merely a dichotomous or trichotomous aggregate of traits.

2. *High authoritarians in comparison with low authoritarians, should select relatively more nonlove-oriented responses for the sex and aggression areas respectively, than for the remaining areas.* This prediction was based on the supposition of authoritarian personality

[1] An abbreviated F scale of 12 items was used; it contained Items 4, 9, 13, 18, 21, 25, 26, 34, 37, 42 from Form 45, Item 44 from Form 60, and Item 55 from Form 78 of the original F scale (1). The TFI items were the 12 comprising the short form presented by Levinson and Huffman (4).

theory that children's sexual and aggressive behavior are especially unacceptable to authoritarian parents, and especially subject to control by discipline.

If nonlove-oriented measures are the characteristic disciplinary techniques of high authoritarian parents, and if children's sexual and aggressive behavior is especially unacceptable to them, then it might be expected that authoritarian parents would select more nonlove-oriented responses for these areas than for the other areas. To test that portion of the prediction pertaining to aggression, each high and low authoritarian mother was given a score based on the differences between the percentages of nonlove-oriented responses in the aggression area and the combined value of the remaining areas (minus sex). The t ratio of 1.8 between the mean scores of the highs and lows, was close to significance at the .05 level.

The failure of the data to attain an acceptable level of significance may reflect the arbitrary grouping of genotypically different items in the particular categories adopted. The behavior situations classed under aggression included items differing markedly from each other in interpersonal significance. Perhaps behaviors embodying threat to authority relationships may be more germane to the authoritarian personality dimension than the particular situations chosen for inquiry. For example, one group of aggressive behaviors that may be interpreted as predominantly involving threat to parental authority discriminated significantly between highs and lows in an individual situation analysis, while the remainder did not. Aggression (and other areas of problem behavior as well) may require finer analysis to yield satisfactory categories in terms of interpersonal significance.

Not only did the data fail to support that part of the prediction pertaining to sex, but as Table 1 indicates, the differences between highs and lows in the selection of nonlove responses were smaller in the sex area than in any other area. It might be speculated that this finding indicates considerable enlightenment in the attitudes of the sample studied towards sexual behavior in their children. The writer was impressed with the casualness and acceptance which the mothers displayed concerning sexual manifestations in their children. Many of the mothers indicated that their own mothers would have been shocked by such behavior, and would not have been so tolerant. While this suggests a cultural trend toward increased leniency in middle-class attitudes toward the expression of sex, it also is very likely in considerable part a reflection of the cultural sophistication of the sample. Moreover, sexuality per se was generally not recognized nor accepted as intrinsic in the particular behavior, but the behavior was generally equated to intellectual curiosity, or otherwise rationalized as part of the childhood learning process.

As may be observed in Table 1, the class of behavior control techniques preferred varied with behavior area. To determine if these variations exceeded chance probability, analyses of variance were computed in which the six behavior areas and the three authoritarianism levels were the independent variables and the classes of disciplinary techniques were the dependent variables. The data summarizing the analysis of variance for nonlove-oriented disciplinary techniques are presented in Table 2. Since love and nonlove response frequencies are essentially mirror images of each other, the results of the analysis of variance for love-oriented discipline were substantially the same. Since the distributions were heterogeneous in variance and form, the F ratios were interpreted as inflated. While the very large F ratio (137) suggests that it would be quite safe to assume that the main effect of behavior areas is significant, the comparatively small F ratio (2.2) for interaction effects does not warrant the conclusion of a significant interaction. This finding corresponds to the nonsignificant results for Prediction 2.

Inspection of Table 1 indicates that nonlove-oriented responses were selected most frequently in the independence and aggression areas, and least frequently in the sex and dependence areas. One possible explanation is that the independence and aggression areas depict behaviors that are more threatening to parental control and authority. However, many of the mothers explained their choice of response to the independence items referring to certain behaviors as potentially dangerous to the child by indicating that threat to their child's safety was the one time they did not hesitate to resort to physical punishment. It might be surmised that ag-

TABLE 2
ANALYSIS OF VARIANCE OF PREFERENCE FOR NONLOVE-ORIENTED DISCIPLINARY TECHNIQUES

Source	*df*	*SS*	*MS*	*F*	*P*
Levels of authoritarianism	2	46796.62	23398.31	33.52	.01
Subjects within levels	123	85864.80	698.09		
Behavior areas	5	217209.28	43441.86	137.07	.01
Level × Area	10	7014.05	701.41	2.21	.05
Area × Subjects within levels	615	194917.34	316.94		
Total	755	551802.09			

gressive behavior on the part of the child might provoke parental aggressive retaliation. The most frequent reason advanced by the mothers for their reluctance to select nonlove-oriented responses for dependency behavior was fear that these techniques might cause the child to feel rejected.

3. *The low authoritarian's description of her child-rearing philosophy and related attitudes should contain more attitude statements with the judged consequence of maintaining the child's approach tendency to her, than should the high authoritarian's.* The high authoritarian's "philosophy" is expected to be judged as more likely than the low authoritarian's to produce and strengthen an avoidance tendency in the child. For the purpose of testing this prediction, each mother was asked the following open-end questions: "What is your philosophy of child-rearing?" and "What do you think are the most important things to take into consideration when dealing with child behavior problems?" To reduce the scope of the task of rating, only the protocols of the mothers with the 30 highest and lowest authoritarianism scores were used in the statistical test. These protocols were submitted to two additional judges who rated them in terms of the anticipated effect the depicted behavior would have on the tendency of the child to approach or avoid the mother. The protocols were in no way identified with regard to authoritarianism and, in fact, the judges were unacquainted with authoritarian theory or the hypotheses of the study. Each protocol was divided into meaningful, coherent idea units on a purely a priori basis, in an attempt to facilitate comparability of ratings. The task of the judges was to sort each unit of the mother's statements into an "approach" category, i.e., behavior thought to strengthen the child's approach tendency; an "avoidance" category; or a "miscellaneous" category. Although the number of units in the protocols varied, the mean number of units of the highs was close to that of the lows. However, as a precaution and in an attempt to compensate for unequal length of protocols, the score for each protocol consisted of the summed approach rating minus the summed avoidance rating. Thus, the final score for each protocol is a discrepancy score and is assumed to represent the relative excess of approach over avoidance (or vice versa).

Fairly high interrater agreement is reflected in the product-moment correlation of .72 for the lows and .73 for the highs. It was concluded that there was sufficient reliability in the judges' ratings to warrant their combination in mean judgments. The mean discrepancy score of the lows was 36.1, *SD*, 7.6; that of the highs 30.5, *SD*, 6.42; each case representing an excess of approach over avoidance judgments. The *t* test computed to determine the significance of the difference between the scores of high and low authoritarians yielded a value of 3.08, $P < .01$, in support of Prediction 3.

The confirmation of this prediction suggests that the structured interview schedule has construct validity (2), and that the socialization techniques differentially preferred by high and low authoritarians have conceptual significance in terms of approach and avoidance principles.

An attempt was also made to determine if maternal child-rearing preferences varied significantly as a function of such variables as sex and age of the child, and if interaction between either of these variables and authoritarianism could account for differences in the preferences. The child's age range was arbitrarily divided into three groups: 2½–3½; 3½–4½; 4½–5½, and three analyses of variance were carried out, in which the three age ranges and three levels of authoritarianism were the independent variables, and the three disciplinary techniques were successively the dependent variables. Disciplinary techniques varied significantly as a function of age only in the case of "ambiguous" responses. There was a slight, but significant, tendency for "ambiguous" responses to be selected increasingly by all three groups, as child's age increased (*F* ratio 3.9, $P < .05$).

In the case of nonlove-oriented discipline, but only then, child-rearing preferences varied significantly as a function of child's sex. All three authoritarian groups responded slightly more often with nonlove-oriented responses when the children were boys, than when they were girls (F ratio 5.9, $P < .05$). As in the case of the age variable, there was no significant interaction between child's sex and authoritarianism.

DISCUSSION

These findings emphasize a conclusion that many have long taken for granted. They show that the choice of behavior control techniques, in a variety of socialization situations, is to a significant extent conditioned by measurable personality factors in the mother. Whiting and Child (7) have argued that parental personality is one critical link in the mediation of culture from generation to generation. The present study, like theirs, was not designed to clarify the antecedents and consequents involved in this process. However, the results do suggest that different emphases or alternative customs of the culture are mediated as a function of maternal personality.

The significance of the present results for authoritarian personality theory depends rather critically on the status of one assumption: that the authoritarian mother's relative preference for nonlove-oriented punishment is a cause as well as an effect of authoritarianism. Assuming that the children of authoritarians themselves tend to be authoritarian, one can speculate about the role of avoidance tendencies in the etiology of authoritarianism. It may be that the child of authoritarian parents is highly ambivalent toward them (and other authorities) because of a heightened avoidance gradient, resulting in more closely balanced approach and avoidance tendencies. His later behavior toward superordinates typically reflects "coping techniques"—such as emphasis on the maintenance of social distance, interaction between formal roles rather than persons as individuals, etc.,—designed to manage this ambivalence. But, as many writers on authoritarian personality have contended, there remains an underlying resentment and hostility toward authority that is revealed in such displaced forms as scapegoating and prejudice.

The speculative tenor of the above remarks could be avoided if the child's actual behavior toward and perception of his parents, and reaction to different forms of socialization pressure, were determined by direct and indirect means in subsequent research.

SUMMARY

The primary purpose of this study was to investigate the way in which the mother's authoritarianism affects her selection of behavior control techniques in a variety of socialization situations. Specific predictions were formulated in an attempt to combine the hypotheses and insights of authoritarian personality theory (1) with the socialization model suggested by Whiting and Child (7). One hundred twenty-six mothers of preschool children responded in a structured interview to 38 behavior situations which commonly arise and call for maternal decisions about mode of control. These items were classified as relating to feeding, cleanliness and toilet training, sex, aggression, dependence, and independence; the preferred response of the mother could in turn be classified as love-oriented, nonlove-oriented or "ambiguous."

The main results of the study may be summarized as follows:

1. As predicted, authoritarians show a consistent tendency to select more nonlove-oriented and fewer love-oriented responses than nonauthoritarians. When the middle-range authoritarians are included, inspection indicates the relationship is linear.

2. There are significant variations in the proportion of love- and nonlove-oriented responses selected as a function of the behavior area involved. Thus, aggression and independence involve behaviors that provoke more nonlove-oriented responses, whereas sex and dependency behaviors call forth fewer nonlove-oriented responses.

3. There is no significant interaction, however, between authoritarianism and behavior area. That is, in terms of the a priori classification of items into six different areas, authoritarianism does not discriminate significantly better in one area than in another. It is quite possible, though, that different item arrangements (in terms more precisely relevant to authoritarianism) would show that the authoritarian's relative preference for nonlove responses is to some extent conditioned by the situation.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper & Bros. 1950.
2. CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, **52**, 281–302.
3. HART, I. Maternal child-rearing practices and authoritarian ideology. Unpublished doctor's dissertation, Duke Univer., 1956.
4. LEVINSON, D. J., & HUFFMAN, PHYLLIS E. Traditional family ideology and its relation to personality. *J. Pers.*, 1955, **23**, 251–273.
5. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1956.
6. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, **47**, 135–234.
7. WHITING, W. M., & CHILD, I. L. *Child training and personality: A cross-cultural study*. New Haven: Yale Univer. Press, 1953.

Received November 12, 1956.

CATHARSIS AND THE RELIEF OF HOSTILITY

PHILIP WORCHEL

University of Texas

CATHARSIS, in one form or another, occupies a central position in all psychotherapeutic procedures. Corsini and Rosenberg (4), in reviewing the dynamics of group therapy, found that ventilation, or the release of feelings and expression of ideas usually repressed in other nontherapeutic situations, was one of the most frequently mentioned mechanisms in about 300 articles examined. Despite its importance, considerable confusion exists concerning the meaning and rationale for the efficacy of catharsis. Rotter (12) pointed out the many different ways in which catharsis is currently viewed and evaluated the different techniques of attaining catharsis from a learning point of view. Dollard and Miller (5) emphasized the importance of the events following the expression of feelings or past memories rather than the sheer act of catharsis. The therapist as a nonpunitive agent helps to extinguish certain avoidance reactions or to develop new learning by pointing out associations not originally perceived by the patient. Alexander and French (1) attributed the therapeutic outcome following catharsis to insight or new learning rather than to any purging effect of the unconscious. New learnings may also occur merely as a result of new relationships perceived by the patient as he verbalizes previous experiences.

Relatively few investigations have been undertaken to validate the clinical impression of the psychotherapeutic efficiency of catharsis or test any hypotheses as to its rationale. Haggard (7), using electric shock to induce disturbance, tested the relative efficacy of "catharsis-information" and experimental extinction or "rest" on reducing the general level of disturbance as measured by autonomic variables. Though he found that "catharsis-information" was significantly more effective, the design did not permit determination of whether "catharsis" or the straightforward "information" that shock would not be given was the essential variable. McKeachie *et al.* (8) found that students who were encouraged to write comments about their questions on an examination made higher scores than students who had conventional answer sheets. The subjects, however, who were instructed to write explanations made slightly, but not significantly, better scores than those who were instructed to write feelings. The results raised doubts concerning the efficacy of "blowing-off steam." Wiener (13) investigated the relative effectiveness of two counseling techniques: reassurance-interpretation and catharsis-reflection. Though the combined counseled groups showed consistent but not always significant improvement in performance in comparison with the noncounseled groups, there were no indications that either of the two experimental counseling techniques was more effective than the other.

These ambiguous and contradictory results concerning the efficacy and the nature of effects of catharsis may be due to the untested assumption that catharsis in whatever form is equally applicable to all kinds of disturbances regardless of the underlying dynamics. It is suggested, however, that catharsis is effective to the extent that it sets up conditions for reducing threat to need-satisfaction patterns. Conflict may arise from a need to express hostility or a threat to self-esteem or between the moral values of the super-ego and the libidinal impulses of the id. Mere expression of anxiety feelings concerning an anticipated failure, for example, may not, in and of itself, eliminate the underlying threat. An analysis of the source of conflict in the expression of hostility suggests certain hypotheses concerning the role of various responses in the reduction of the conflict. The goal of the hostility drive is attack and removal of the stimulus arousing hostility. The threat to such direct behavior, however, is the anticipated punishment derived from earlier experiences in expressing hostility in this manner. It is hypothesized, therefore, that hostility will be reduced if the therapeutic situation provides socially acceptable conditions for the expression of hostility or attack at the instigator.

There has been some question (12), however,

concerning the generalization of the relief of such feelings when expression of hostility is made to a third party (e.g., the therapist) or displaced upon some object or person other than the original instigator. On the one hand, expression directly to the instigator, even under nonpunitive conditions, may still arouse anxiety concerning the anticipated consequences of such attack. On the other hand, a number of experiments (9, 10) have indicated the generalization effect occurring under conditions of displacement. A second objective, therefore, of the present series of studies is to compare the effect of expressing hostility directly to the instigator, and expressing hostility indirectly to or about others.

METHOD

Subjects

From nine sophomore and junior classes in elementary and educational psychology, 52 subjects (*Ss*) were selected from each of the three largest groups, 52 *Ss* from two smaller groups, and 52 *Ss* from the two smallest groups. The five groups were randomly assigned to the experimental treatments and were identified as follows: Group I, Catharsis-Experimenter (E); Group II, Catharsis-Assistant (A); Group III Catharsis-Fantasy (F); Group IV, Control (Talk); Group V, Control (Silence). Sex was about evenly divided in all groups: in Group I, there were 23 women and 29 men; in Group II, 25 women and 27 men; in Group III, 23 women and 29 men; in Group IV, 28 women and 24 men; and in Group V, 22 women and 30 men.

Experimental Tasks

Basically, the design involved the arousal of hostility in groups of *Ss*, exposing them to different conditions of ventilating feelings, and then testing the efficiency of the experimental conditions in the reduction of the disruptive effects of stress on performance. Hostility was aroused by an insult technique similar to the one used by Feshbach (6) with one modification; insult was directed at the performance of *Ss* in an ego-involved task, an intelligence test. In the present experiment, catharsis refers to the release of hostility feelings, verbally expressed by volunteers in a group, under nonpunitive conditions. The tasks used to assess the degree of the reduction of stress were a digit-symbol test and a test of incidental recall of the titles of a bogus intelligence test. The basic assumption underlying the use of performance measures as indices of the disruptive effects of stress follows from the interference interpretation of the effects of frustration (2) by Child and Waterhouse (3): Responses such as anxiety and anger evoked by frustration in one activity may interfere with the quality of performance in a second activity.

Procedure

To increase involvement, *E* introduced himself as professor of psychology and director of clinical psychology. Each group of *Ss* was told that as a part of the general routine of the Department of Psychology, students were given a series of tests measuring learning ability and intelligence. The scores were entered in each student's file for purposes of guidance, recommendations for scholarships and jobs, and consideration for admission to graduate study if the students applied for such admission.

As a pre-experimental test of learning ability, each group was given a one-minute digit-symbol test, the code consisting of letters and geometrical figures. Then a bogus intelligence test (14) entitled "The University of Texas Intelligence Test", and consisting of 14 subtests, each prominently labeled by a single word describing the contents, was administered to each group. The instructions, read verbally by *E*, indicated that this test consisted of a number of subtests, each highly related to academic success, and that the time limits of 45 to 90 seconds allowed sufficient time for most students to complete each subtest. At the beginning of each subtest, *E* said, "This test is called a test of _____. It is a measure of your ability to _____. Are you ready? Begin!" The time limits for the first two subtests permitted most of the students to complete them, but for the remaining 12 subtests, the material became much more difficult and the time for completion was sufficient for only a very few *Ss*. In order to arouse hostility feelings, *E* went around the room, critically remarking at the poor and slow performance of the *Ss*, urging more rapid solutions, and stating how much faster and better other classes were in this test.

The conditions during the period of eight minutes following the intelligence test varied for each group depending upon the experimental situations involved. In Group I (Catharsis-E), *Ss* were informed that there would be a slight rest period before the next test and that during this time the examiner would like to know how they felt about the test and the administration of the test. For Group II (Catharsis-A), *E* said that he would have to go to his office to bring more test material but that a teaching assistant would come in to take over until his return. The assistant, who had been instructed beforehand to profess ignorance concerning the events in the classroom, asked the group to inform him as to what had happened and to give him their reactions. The role of *E* or assistant for each group was merely to accept the statements without evaluation or comment except for a "hmm" or repetition of the statement. At the end of eight minutes, *E* walked in with more papers and continued the experiment. During the rest period for Group III (Catharsis-F), *E* mentioned that the university had surveyed student opinion a year ago concerning "irritations at the campus." As a follow-up on the previous survey, the class was asked to discuss their evaluation of the present situation.

In order to control the effect of "just talking" as against actually expressing hostility, a "neutral" topic was assigned for discussion in Group IV (Control-Talk). There had been recent publicity concerning the

shortage of science majors. The students were asked to suggest methods of encouraging more students in high schools to major in science. Finally, *E* lectured to Group V (Control-Silence) for eight minutes on the preparation of majors in psychology, allowing neither "just talking" nor the expression of hostility.

Following the "rest" period, each group was given another form of the digit-symbol test (numbers and different symbols). At the end of one minute, *Ss* were asked to write down the titles or brief descriptions of all the subtests in the intelligence test (incidental recall). All groups were tested within a two-day period to minimize communication to other *Ss*.

When all five groups had been tested, *E* described the nature of the experiment and full procedure with particular emphasis on the function of the bogus intelligence test to all *Ss* in order to relieve any doubt concerning their intellectual status.

RESULTS

Expressions of Hostility

A graduate assistant, unaware of the objectives of the present experiment, sat with the *Ss* in each group and made notes of their remarks during the experimental session. Though it is difficult to assign a global rating of the intensity of expression for the group as a whole, the nature of the remarks indicated rather clearly that reports of expressed hostility were more frequent and intense in Groups II and III than in Group I. Even though *Ss* in Group I were encouraged to express their feelings, they gave mild criticisms only of the test. They suggested that the time limits were too short, *E* did not give sufficient time for them to read instructions, the items were too difficult, and the test was not a valid measure of their intelligence. One might infer that hostility, if present, was directed toward the test. In Group II, *Ss* talked more directly about *E* to the assistant. They said he "talked too much," "he was too cocky," "who did he think he was," "what the hell was his name," "if he kept his mouth shut they could have done better." Group III was very critical of staff members and teaching assistants: "Professors rewrote their books every year so that students have to buy new textbooks," "professors were often unprepared for their lectures," "some had no business to teach," "why don't professors stay in their offices to see students," "teaching assistants gave unfair exams and didn't know any more than the students." *Ss* seemed able to report more hostile feelings about the *E* to somebody else than to the *E* directly. When they had to express their feelings directly to *E* in Group III, however, they expressed more hostility about other authority figures (professors and assistants) than about *E* himself. In the control situation of Group IV, *Ss* suggested methods of encouraging more science majors such as increased publicity, more use of audiovisual techniques in science classes, and emphasis upon applications of science courses. In addition, however, there were a few critical remarks on "poor science teachers" and "inadequate pay for good teachers."

TABLE 1

ADJUSTED POST-TREATMENT LEARNING SCORES ON THE DIGIT-SYMBOL TEST FOR THE FIVE GROUPS

Group	Observed Mean $\bar{Y}_i$	Observed Mean $\bar{X}_i$	Adjusted Mean $(\bar{Y})$
I Catharsis-E	44.96	28.44	44.52
II Catharsis-A	46.15	28.90	45.34
III Catharsis-F	44.48	27.69	44.65
IV Control (talk)	42.61	28.17	42.37
V Control (silence)	38.83	26.11	40.24

Learning

It was predicted that hostility feelings would be reduced under conditions encouraging expression of attack or hostility. Specifically, therefore, scores on the digit-symbol test should be higher for the three experimental than for the two control conditions.

For the digit-symbol test, Table 1 shows the pre-experimental and postexperimental means for all five groups. Since the pre-experimental means differed in the five groups, an analysis of covariance was applied. The adjusted mean scores are shown in the last column.

Since an obtained F ratio of 11.43 for the treatment variable was significant well beyond the .01 level, tests of significance among the five groups of *Ss* were performed and are reported in Table 2. Since adjusted mean digit-symbol scores for Groups I, II, and III are significantly higher (.01 level) than those for Groups IV and V, the hypothesis seems clearly supported. There is no significant difference, however, among the adjusted means of Groups I, II, and III. Nevertheless, Group IV, which merely discussed a neutral topic,

TABLE 2

TESTS OF SIGNIFICANCE BETWEEN THE ADJUSTED MEANS OF THE DIGIT-SYMBOL SCORES FOR THE FIVE GROUPS

Groups	Adjusted Mean Diff.	*t* Ratio	*P*
I and II	.82	.96	—
I and III	.13	.15	—
I and IV	2.15	2.43	.01–.02
I and V	4.28	4.84	<.01
II and III	.69	.78	—
II and IV	2.97	3.36	<.01
II and V	5.10	5.76	<.01
III and IV	2.28	2.58	<.01
III and V	4.41	4.98	<.01
IV and V	2.13	2.41	.01–.02

TABLE 3

TESTS OF SIGNIFICANCE BETWEEN THE MEANS OF THE FIVE GROUPS ON THE POST-TREATMENT TEST OF INCIDENTAL RECALL

Groups	$\overline{X}_D$	*t* Ratio	*P*
I and II	1.21	3.36	<.01
I and III	.77	2.14	.02–.05
I and IV	.68	1.95	.05
I and V	.85	2.36	.01–.02
II and III	.44	1.22	—
II and IV	1.89	5.22	<.01
II and V	2.06	5.72	<.01
III and IV	1.45	4.00	<.01
III and V	1.62	4.50	<.01
IV and V	.17	.50	—

did significantly better (.01–.02 level) than Group V, where the *S*s maintained silence.

Incidental Recall

The results of the test for incidental recall not only confirm the findings on the digit-symbol test of significantly superior performance by the three experimental over the two control groups, but also point to a significantly superior performance by Groups II and III over Group I. An obtained *F* ratio of 12.48 was significant beyond the .01 level for the treatment variable. As shown in Table 3, the mean incidental recall scores (number of titles recalled) of 6.50 and 6.06 for Groups II and III, respectively, are significantly higher than the mean of 5.29 for Group I. All three mean scores for the experimental groups are significantly higher than those of 4.61 and 4.44 for the two control groups (Groups IV and V, respectively). No difference between the mean scores of Groups II and III or between the two control groups was demonstrated. Thus, incidental recall following expression of hostility to the instigator about other authority figures or following expression of hostility about the instigator to another authority figure seems to have a significantly more facilitative effect on performance than the direct expression of hostility to the instigator.

DISCUSSION

If the learning and incidental recall indices can be accepted as indirect indicators of the effects of the relief of aggressive tension, then the results clearly indicate (*a*) the therapeutic effects of catharsis on the relief of aggressive tension, and (*b*) the differential effects of varying cathartic conditions. It may be that hostility was relieved by ventilation or that expression of hostility about *E* or other similar figures in the cathartic situation permitted a reduction of hostility feelings towards *E* in the postexperimental test situation (generalization). Certain limitations in the present experiment, however, should be recognized. Since groups rather than individuals were employed in the study, relatively few *S*s reported their feelings during the brief cathartic period. It may be that the silent members relieved hostility feelings vicariously as others expressed aggression, or the effect may have been largely restricted to the more verbal members. The experiment warrants replication so that the differential effects on the verbal and silent members may be studied.

Another factor to be considered in interpreting the results is the brief and "single-shot" nature of the cathartic period. It is possible that hostility would be more extensively relieved in Group I (direct expression to *E*) if a number of cathartic sessions were used. A number of trials might be necessary for *S*s in Group I before the free expression of hostility feelings could occur. Though the results indicate that hostility feelings were reduced as much if not more in Groups II and III than in I, nothing can be said as to the duration or generalization of these feelings (postexperimental sessions). Experiments on displacement (9, 10, 11) suggest that the approach gradient would rise for the original goal after approach to similar stimuli. Brief encounters with *S*s in the corridors after the testing of the initial groups and

prior to the explanation of the experiment suggest that hostility was still strong towards *E*. In the present experiment, the conditions in Groups II and III might be conceived as involving displacement of hostility feelings. In one case, the object of attack, *E*, was kept constant, but the person to whom the feelings were expressed varied from *E* (Group I) to the assistant (Group II). In the other case, the feelings were expressed to the same person, *E*, but the object of attack varied from *E* (Group I) to other staff members (Group III). Traditionally, the latter case is more appropriately an example of displacement since the object of attack is similar but not identical to the original instigator. The former case, more like the therapeutic situation, may nevertheless be considered an example of displacement since the punitive agent to whom hostility is expressed differs from the original.

One final consideration involves the role of *E* in the two control groups. *E* had to be present in order to equate the condition of his presence in Groups I and II, but a more complete design to test the effect of the assistant in Group II should include a replication of the two control groups with the assistant rather than *E* in charge of the eight-minute control period. It is possible that the lecture by *E* in Group V might have heightened temporarily any feelings of hostility, or that discussion concerning increasing the number of science majors in Group IV may have reduced hostility feelings through some critical comments about poor teaching. Either situation, the lecture by *E* in Group V or the discussion of science majors in Group IV, might explain the significantly better results for Group IV on the digit-symbol test.

SUMMARY

The purpose of the present study was to study the effect of different cathartic techniques on the relief of hostility feelings. From nine sophomore and junior classes in elementary and educational psychology, 52 *Ss* were selected from each of the three largest groups, 52 *Ss* from two smaller groups, and 52 *Ss* from the two smallest groups. All groups were given a bogus intelligence test consisting of 14 subtests to produce ego-involvement. Hostility was aroused by the use of "insult" during the administration of the tests. During the "cathartic" sessions of eight minutes, Group I was asked to express criticism of the administration of the test. For Group II, *E* mentioned that he was going to his office for other test material and that an assistant would take over for the next few minutes. The assistant professed ignorance of the proceedings and asked *Ss* to describe the previous events and state their reactions to them. Group III was requested to give a follow-up evaluation of a study made a year ago on campus irritations. Group IV, a control group, was asked to discuss a "neutral" topic, and V, another control group, was not permitted to say anything while *E* lectured on the preparation of majors for psychology. A pre- and post-treatment digit-symbol test and an incidental recall test of the 14 titles of the subtests of the intelligence test were used as measures of the effect of the relief of hostility feelings. The results were as follows:

1. *Ss* in Groups II and III were able to express more hostile feelings than those in Group I, who were asked to express feelings about *E* directly to him.

2. Adjusted digit-symbol scores were significantly higher for the three "cathartic" groups (I, II, III) than for the two control groups (IV, V).

3. There were no significant differences in the adjusted mean digit-symbol scores among the three experimental groups.

4. On incidental recall, *Ss* in Groups II and III were significantly higher than all other groups.

5. *Ss* in Group I gave significantly higher incidental recall scores than those in the control groups (IV and V).

REFERENCES

1. ALEXANDER, F., & FRENCH, T. M. *Psychoanalytic therapy: Principles and application*. New York: Ronald Press, 1946.
2. BARKER, R. G., DEMBO, TAMARA, & LEWIN, K. Frustration and regression: An experiment with young children. *Univer. Iowa Stud. Child Welf.*, 1941, **18**, 1–314.
3. CHILD, I. L., & WATERHOUSE, J. K. Frustration and the quality of performance: I. A critique of the Barker, Dembo, and Lewin experiment. *Psychol. Rev.*, 1952, **59**, 351–362.
4. CORSINI, R. J., & ROSENBERG, B. Mechanisms of group psychotherapy: Processes and dynamics. *J. abnorm. soc. Psychol.*, 1955, **51**, 406–411.
5. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.

6. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, **50,** 3–11.
7. HAGGARD, E. A. Some conditions determining adjustment during and readjustment following experimentally induced stress. In S. S. Tomkins (Ed.), *Contemporary psychopathology.* Cambridge, Mass.: Harvard Univer. Press, 1947. Pp. 529–544.
8. MCKEACHIE, W. J., POLLIE, D., & SPEISMAN, J. Relieving anxiety in classroom examinations. *J. abnorm. soc. Psychol.*, 1955, **50,** 93–98.
9. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43,** 155–178.
10. MURRAY, E. J., & BERKUN, M. M. Displacement as a function of conflict. *J. abnorm. soc. Psychol.*, 1955, **51,** 47–56.
11. MURRAY, E. J. A case study in a behavioral analysis of psychotherapy. *J. abnorm. soc. Psychol.*, 1954, **49,** 305–310.
12. ROTTER, J. B. *Social learning and clinical psychology.* New York: Prentice-Hall, 1954.
13. WIENER, M. The effects of two experimental counseling techniques on performance impaired by induced stress. *J. abnorm. soc. Psychol.*, 1955, **51,** 565–572.
14. WORCHEL, P. Anxiety and repression. *J. abnorm. soc. Psychol.*, 1955, **50,** 201–205.

Received November 30, 1956.

ASSIMILATION AND CONTRAST EFFECTS IN REACTIONS TO COMMUNICATION AND ATTITUDE CHANGE[1]

CARL I. HOVLAND
Yale University

O. J. HARVEY
Vanderbilt University

AND

MUZAFER SHERIF
University of Oklahoma

THIS paper presents an experiment on reactions to communication and on attitude changes by individuals whose initial stands on a controversial social issue diverged in varying degrees from positions advocated in communication. Study of the relationship between subject's (*S*'s) attitude and the position advocated in communication may help resolve some apparently contradictory effects of communication aimed at changing attitudes.

Attempts to change attitudes in the direction advocated by communication on a social issue at times produce shifts in the direction opposite to that intended—the "boomerang effect." While numerous investigators have reported shifts of average test scores in the direction of communication (4), a fairly common finding, even in these studies, is that some individuals shift their stand *away* from that presented in communication. Several studies reporting both positive and negative shifts in attitudes toward out-groups following communication are summarized by Williams (20).

Thus, at times, persuasive communication produces a bi-modal distribution of attitude scores (11, pp. 874–875). For example, Remmers (12) obtained positive shifts on average scores following communications on conservation, social insurance and labor unions, but the latter communication "sharply divided the group into two opposing tents" (12, p. 201). In Knower's study prior to repeal of prohibition, "wet" communication to generally "dry" *S*s and "dry" communication to generally "wet" *S*s resulted in shifts in both positive and negative directions in each group (8). A rather striking instance of such opposing effects was reported by Wilke (18), whose antiwar communication was presented at a time when the student population from which *S*s were drawn was divided in controversy over this very issue.

A few studies, such as those by Manske (10) and Russell and Robertson (14), have found group shifts in the direction opposite to the stand presented in communication. A related finding reported by Williams and Remmers (19) following communication on a rural issue was reduced variability and a comparatively less favorable stand by a group of rural youth in contrast to the increased variability in an urban group. Some authors who have obtained results in the direction away from communication insert the suggestion that too great divergence between *S*'s stand on the issue and the stand presented in communication may have been responsible.

On the other hand, data from several recent studies suggest that the extent of influence increases as a function of the distance between position of communication and position of the recipient. Goldberg (2) reports that the greater the discrepancy between an announced group norm and the *S*'s own judgment the greater the change produced. In a still unpublished paper, French and Gyr (summarized in 1) found a positive correlation between the degree of deviation between inducer and inductee and the amount of change. In a perceptual task, Lubin and Fisher (9) obtained an increase in conformity to a partner's judgment as the distance increased up to a point where with great distance the proportion of movement declined. Hovland and Pritzker (6) have shown that the larger the change in opinion advocated the greater the amount of

[1] This investigation was conducted as part of the Yale Communication Research Program. This series of studies, devoted to an analysis of factors related to attitude and opinion formation, is financed by a grant to Carl I. Hovland from the Rockefeller Foundation, whose support is gratefully acknowledged. Thanks are extended to Dr. Charles Shedd, now of Berea College, Mr. William Smith, and Mr. Richard Disney who participated in developing the procedures and in writing the communications used. We are also grateful to Dr. Laurence H. Snyder, Dean of the Graduate College, University of Oklahoma, for necessary administrative arrangements, and to Drs. Irving L. Janis, Robert P. Abelson, and Jack W. Brehm for suggestions made in the course of reading the manuscript.

change produced. The issues that were employed did not pertain to strongly rooted attitudes, but represented opinions on rather factual topics. The communicators used were authorities respected by the recipients.

A comprehensive analysis may be necessary to handle the above findings as well as results of studies employing communication on controversial social issues which suggest that too great distance between *S*'s attitude and the position advocated in communication produces "boomerang effects." One possible approach is through analysis of processes underlying judgments of motivationally neutral material, exemplified by psychophysical stimuli, as well as ego-involving verbal material, exemplified by controversial social issues. Investigations of judgmental processes and their theoretical relevance for the study of reaction to communication and attitude change will be presented in a forthcoming volume.

One set of findings suggests that *S*'s position shifts toward the stand advocated in communication when the topic is not highly ego-involving and *S*'s position does not diverge in the extreme from the stand advocated. This result seems akin to extension of an established reference scale in judgment of weights or inclinations following introduction of anchoring stimuli near the end stimulus of the series, as reported by Rogers (13), Heintz (3), and others. Sherif, Taub, and Hovland (16) demonstrated this extension toward the anchor ("assimilation effect") as well as the effect of more remote anchors beyond either end of the stimulus series in constricting the reference scale of judgment and displacing judgment away from the anchor ("contrast effect"). Contrast phenomena may also appear in reaction to communication on an issue which is not highly involving for *S* if the position advocated is removed sufficiently from *S*'s position, as the Lubin and Fisher results seem to indicate.

The above analysis based on findings from judgment studies suggested the further possibility that when *S*s have established attitudes and are personally involved in a controversial social issue, their "own stand" functions as the major anchorage affecting reaction to and evaluation of communication. In this case, communication near *S*'s stand would be assimilated to it, while communication at variance with *S*'s own stand would be displaced still further away ("contrast effect"). Whether assimilation or contrast effects appear would be a function of the relative distance between *S*'s own stand and the position of communication.

Accordingly, in the present experiment communications representing two opposite extremes and one moderate position on an ego-involving issue were presented to *S*s whose initial stands on the issue ranged from one extreme to the other. The following hypotheses were formulated in terms of the effect of the relative distance between *S*'s own stand and the position advocated upon evaluation and placement of communication as well as acceptance-rejection of that position:

1. Reactions to a communication will decrease in favorableness as the distance between *S*'s own stand and the position advocated in the communication increases.

2. In evaluations by *S* of what position is advocated by a communication, the greater the distance between *S*'s own stand and the position advocated in the communication, the greater the displacement *away* from *S*'s position ("contrast effect"). When only a small discrepancy in position exists there will be a tendency for displacement *toward* *S*'s stand ("assimilation effect").

3. With small distances between the position of the communication and that of the *S*, changes in *S*'s opinion in the direction advocated by the communication will occur. With large distances between the stands taken by communication and by *S*, opinion changes in the direction advocated will be infrequent.

METHOD

Issue and Communications

The problem required that a controversial issue be selected, that *S*s' stands on the controversy be ascertained, and that communications advocating various positions be presented to *S*s with differing stands on the issue.

The issue chosen was the controversy over prohibition and repeal in a "dry" state. Shortly before the study began, a referendum was held to determine the fate of existing prohibition laws. The vote favored prohibition by a narrow margin.

In order to differentiate existing stands on the issue, representative statements made during the referendum campaign were collected from leading newspapers in two large cities. In addition, statements were obtained from 500 people in several localities on a random basis. These statements from public and private sources were sorted by 20 judges to secure clearly differentiated

stands actually taken in the state. As a result, eight statements were chosen representing prevailing stands ranging from strong advocacy of prohibition to strong advocacy of repeal. One additional "wet" statement was added as a logical counterpart of the most extreme "dry" stand, giving a total of nine statements. These statements are as follows, (I) being the additional item:

(A) Since alcohol is the curse of mankind, the sale and use of alcohol, including light beer, should be completely abolished.

(B) Since alcohol is the main cause of corruption in public life, lawlessness, and immoral acts, its sale and use should be prohibited.

(C) Since it is hard to stop at a reasonable moderation point in the use of alcohol, it is safer to discourage its use.

(D) Alcohol should not be sold or used except as a remedy for snake bites, cramps, colds, fainting, and other aches and pains.

(E) The arguments in favor and against the sale and use of alcohol are nearly equal.

(F) The sale of alcohol should be so regulated that it is available in limited quantities for special occasions.

(G) The sale and use of alcohol should be permitted with proper state controls, so that the revenue from taxation may be used for the betterment of schools, highways, and other state institutions.

(H) Since prohibition is a major cause of corruption in public life, lawlessness, immoral acts, and juvenile delinquency, the sale and use of alcohol should be legalized.

(I) It has become evident that man cannot get along without alcohol; therefore, there should be no restriction whatsoever on its sale and use.

Three communications of equal length were prepared, each requiring approximately 15 minutes for delivery. Arguments were those actually made by prohibition and repeal advocates during the referendum campaign. Arguments in the three communications were arranged in the same order, but from the viewpoint of the different parties to the controversy. One communication presented an extreme "dry" stand; one an extreme "wet" stand; and one a moderately wet stand, as typified in statement (F) above. All communications were presented by tape recording. The same voice was used in recording wet and dry communications.

Procedure

In the first session, data on *Ss*' attitudes were obtained. At the time, *Ss* were not told that an additional session would be held. A schedule on "public issues" was presented with assurance of anonymity to *Ss*. Following a "dummy" issue (college football) *Ss* responded to the nine representative statements on prohibition. The following instructions were printed on the schedule and read aloud by the experimenter (*E*):

> Below are some statements recently made concerning the wet-dry issue in this region.
>
> Please read *all* of the statements carefully first before making any marks on this page.
>
> Now that you have carefully read all the statements, ***underline*** the ***one*** statement that comes closest to your own point of view on the topic.
>
> There may be other statement or statements which you find not objectionable from your point of view. Put a circle around the letter in front of such a statement or statements which are ***not objectionable*** to you.
>
> Now cross out that one statement which is ***most objectionable*** from your point of view.
>
> There may be other statement or statements which you find ***objectionable*** from your point of view. Cross out the letter in front of such a statement or statements which are ***objectionable*** to you.

These procedures yielded data on *S*'s stand and also on the range of his tolerance for other stands in the series (*latitude of acceptance*), the range of stands he rejected (*latitude of rejection*) and those stands which he did not consider either acceptable or unacceptable. Thus the procedures gave *S* an opportunity not to take a stand in relation to items which he did not include in his latitude of acceptance or latitude of rejection, rather than requiring artificial "indifference" or "neutral" checkings.

From 1–3 weeks after the first session, a communication was presented by tape recording. Arrangements for its presentation were made through a member of the group being exposed and it was introduced as a talk actually made by a proponent of the stand advocated. The wet (repeal) communication was presented to extreme dry *Ss* and unselected *Ss*. The dry (prohibition) communication was presented to extreme wet and unselected *Ss*. The moderate communication was presented to wet, dry, and unselected *Ss*. The *Ss* participated in both "before" and "after" sessions in small groups of 10–30 under close supervision.

Following the communication, the same questionnaire for securing *S*'s stand, latitudes of acceptance and rejection was filled in a second time. Reactions to the communication presented were obtained through scales on like-dislike, reasonable-unreasonable, biased-unbiased, propaganda-fact dimensions. The *Ss* in the moderate communication groups also checked on a graphic rating scale ranging from extreme dry (A) to extreme wet (I) positions what they thought to be the stand taken in the communication itself.

Subjects

Since the objective was to secure *Ss* who were definitely ego-involved, a special point was made of obtaining *Ss* with established and publicly committed dry or wet stands as a validity check of the paper-and-pencil checkings. It was not difficult to select *Ss* in the dry groups on the basis of known information concerning their stand. Two small samples from Women's Christian Temperance Union groups were obtained and a group of Salvation Army workers. The four remaining dry groups were students in preparation for the ministry or in strict denominational colleges. A total of 183 dry *Ss* participated in both sessions. It was much more difficult to obtain *Ss* whose position was known to be wet. However, on the basis of cases personally known to the *Es* or their assistants, 25 wet *Ss* were secured. For comparison, 290 additional *Ss* were obtained representing more moderate positions

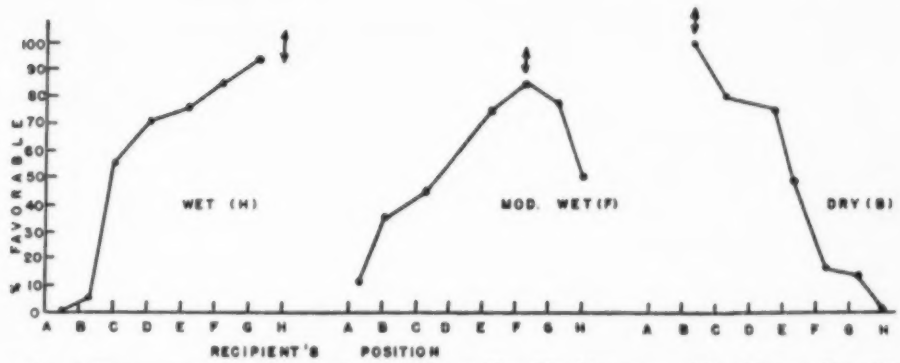

FIG. 1. PERCENTAGE OF FAVORABLE EVALUATIONS OF WET (H), MODERATELY WET (F), AND DRY (B) COMMUNICATIONS FOR *Ss* HOLDING VARIOUS POSITIONS ON PROHIBITION (BASED ON MEAN ACCEPTABLE STATEMENT)

(Positions of communications indicated by arrow)

on the issue. These were college students secured from classes in journalism, speech, education, chemistry, etc. All *Ss* were residents of the dry state where prohibition was a lively topic of controversy at the time of the study. For the reason indicated, it was not possible to match the age levels of *Ss* in the three categories.

RESULTS

Evaluation of the Communication

A five-item scale was used to measure the audience's evaluation of the fairness and impartiality of the communication. In Figure 1, the percentages of *Ss* in each group whose reactions were favorable are graphed for the three communications. On the abscissa the stands of the recipient (mean of statements checked acceptable) are presented and on the ordinate the percentage of favorable evaluations. It will be observed that there is an extremely close relationship between the individual's own stand on the issue and his evaluation of communication. The two communications advocating extreme positions have their peak of favorable responses among those holding corresponding extreme positions. The maximum favorable reaction for the moderate communication is found among those holding a moderate position. The data provide quantitative information to support the expectation from earlier studies (5, 17) that individuals who are in favor of the opinion advocated will consider the communication fair and unbiased, but that those with an opposed stand will regard an identical communication as propagandistic and unfair.

Placement of the Position of the Communication

The principal results of the experiment are those pertaining to the recipient's perception of the stand advocated in the communication. We asked *S* to indicate on a graphic scale ranging from extreme dry position to extreme wet position, what he thought to be the position of the moderately wet communication (at F). From our previous study on assimilation-contrast effects with psychophysical data we predicted that positions differing only slightly from one's own would be "assimilated," while larger differences between one's own position and that of the communication would be exaggerated, showing a "contrast" effect. In Figure 2, results concerning this prediction are presented. The *S*'s own position is indicated on the abscissa and along the ordinate the average rating of the position of the communication is given for the speech in which a moderately wet position (about F) was advocated. The dots indicate the mean placement of the position of the communication for *Ss* who indicate each particular position as their own stand. The squares represent the mean placement of the position advocated when *S*'s position is estimated from the mean of the acceptable positions checked. The dotted line indicates a hypothetical relationship in which *Ss* holding the same position as the communication report its position accurately (at F), those a small distance removed assimilate it to their own position, and those still further removed exaggerate the position revealing a contrast effect.

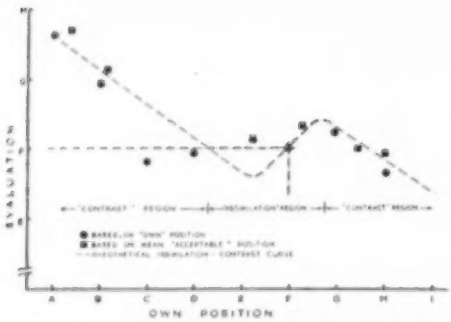

FIG. 2. AVERAGE PLACEMENTS OF POSITION OF MODERATELY WET COMMUNICATION (F) BY *S*s HOLDING VARIOUS POSITIONS ON THE ISSUE, PLOTTED AGAINST HYPOTHETICAL ASSIMILATION-CONTRAST CURVE

The expectation is fulfilled that those at the wet end judge that the communication advocated a drier position than it did and those at the dry end judge that it was advocating a wetter position than it did. Those nearer the position of the communication reported it more objectively. The evidence concerning assimilation deviates somewhat from the theoretical curve. There were too few *S*s with intermediate positions to enable us to determine these positions adequately.

Acceptance-Rejection of Stands

Each *S* was asked to indicate not only the stand which best represented his, but also other stands which were more or less acceptable (i.e., his latitude of acceptance), the stand most clearly opposed to his position, and others not acceptable (i.e., his latitude of rejection). From these data, we find that *S*s with more extreme positions tend to reject more frequently positions not within their latitude of acceptance. The *S*s with middle positions are more apt to rate positions removed from their own as indifferent. In Table 1, data are presented showing the mean number of items rated "also acceptable," "unacceptable" and "neither acceptable nor unacceptable" by *S*s with strong positions (A, B, G, H, I) as compared with those by *S*s with more moderate positions (C, D, E, F). It will be observed that a significantly greater number of items are judged unacceptable by *S*s with extreme stands. Smaller differences between the groups exist in the number rated acceptable. These results tie in with our earlier studies (7, 15) in which we found that *S*s with extreme positions and considerable ego-involvement in an issue have raised thresholds of acceptance and lowered thresholds of rejection in placing items concerning that issue, resulting in smaller latitudes of acceptance and greater latitudes of rejection than for *S*s with moderate positions. These findings can be used in constructing the expected latitudes of acceptance and rejection of the individual for an ego-involving issue once his own position is ascertained and possibly in predicting evaluation of communication which stands at a given distance from his own position. An attempt in this direction is presented in the discussion.

TABLE 1

ACCEPTABILITY OF STATEMENTS IN RELATION TO EXTREMITY OF SUBJECTS' POSITION ON ISSUE

S's Positions	*N*	Mean Number of Items Acceptable	Mean Number of Items Not Checked	Mean Number of Items Rejected
1. Extreme (A, B, G, H, I)	193	2.81	1.48	4.71
2. Intermediate (C, D, E, F)	37	3.05	2.24	3.70
p				<.03

Changes in Opinion for Groups with Different Stands on the Issue

Finally, it would be expected on the basis of the considerations previously discussed that those holding a position at great variance with that being presented in the communication would be relatively little modified in their opinion. The index of opinion change was based on the mean position of the statements checked by *S*s as being "acceptable" to them. A value of 1 indicates the extreme dry end in which only statement A was checked. Correspondingly, a value of 9 is at the wet end with only statement I checked. The positions of the three speeches at B, F, and H would thus be represented by values of 2, 6, and 8, respectively.

Table 2 gives the mean position for five criterion groups before and after the communication, together with the average amount of change. Only the unselected group exposed to the wet communication should be expected to show sizeable change. The four other groups all have initial positions quite removed from the position being advocated. While the

TABLE 2

OPINION CHANGE

Changes in Mean Acceptable Statement for *Ss* with Differing Initial Stands

Group	*N*	Before Comm.	After Comm.	Change in Direction of Comm.
Wet Communication (H)				
Drys	69	2.39	2.34	−.05*
Unselected	92	5.10	5.65	+.55*
Dry Communication (B)				
Wets	25	6.70	6.74	−.04
Unselected	87	5.90	5.78	+.14
Moderately Wet Communication (F)				
Drys	114	2.17	2.26	+.09

* Difference between changes: $p = <.03$ (one tail).

TABLE 3

OPINION CHANGE

Percentage of *Ss* Changing in Direction of Communication or in Opposed Direction

Group	*N*	Change in Direction of Comm.	No Change	Change in Direction Opposed to Comm.	Net Change
Wet Communication (H)					
		%	%	%	%
Drys	69	27.5	49.3	23.2	+4.5*
Unselected	92	52.2	23.9	23.9	+28.3*
Dry Communication (B)					
Wets	25	24.0	56.0	20.0	+4.0
Unselected	87	40.2	33.4	26.4	+13.8
Moderately Wet Communication (F)					
Drys	114	31.6	49.1	19.3	+12.3

* Difference between changes: $p = <.04$ (one tail).

dry and unselected groups exposed to the wet communication do differ substantially in their initial positions, the positions of the wet and unselected groups are similar, since the unselected group (college students) initially holds a rather wet position.

As predicted, there is a significantly greater change for the unselected group given the wet communication than for the dry group ($p = <.03$). As in all studies in which the variable with which we are concerned cannot be directly manipulated, the possibility must be considered that there are factors other than position on the issue which are correlated with it and provide the basis for the relationship obtained.

In Table 3, results are presented for the same groups in terms of the percentage of *Ss* who are influenced either in the direction advocated by the communication or in the direction opposed to it. The net change column represents the differences between those changed in the direction advocated and those changed in the opposite direction. The net change is greater among the unselected group than among the drys when a wet communication is presented ($p = <.04$). Twenty-eight per cent of the former and only four per cent of the latter changed in the direction advocated. It will also be seen that the amount of change produced among the drys by a moderately wet communication (at F) is greater than that produced by the more extreme wet communication (at H). The net proportion changed by the former is 12 per cent, while it is only 4 per cent for the latter. This difference, however, is not statistically significant and, in any case, it is difficult to establish that the two different communications were inherently equivalent in persuasiveness.

It will be noted that by confining our comparisons to groups who were initially selected as having known differences in their stand on the issue (either wet or dry) we were able to avoid the regression effects often found in studies where amounts of change are compared for groups differentiated solely on the basis of initial scores on the attitude measure. Under the latter conditions, those with more extreme scores often regress toward the mean and this can either obscure the effect or produce a relationship artifactually.

Finally, it will be noted from Table 3 that the predominant response among *Ss* holding extreme positions is to remain uninfluenced by the communication. Even where *Ss* in these groups were influenced, changes were seldom found in the item they checked as best representing their own position. Instead, there was typically only an increase in the number of other statements intermediate between their own position and that of the communication which they would check as "also acceptable."

DISCUSSION

The results presented lend support to the three hypotheses pertaining to evaluation and placement of communication and attitude

TABLE 4

HYPOTHETICAL LATITUDES OF ACCEPTANCE AND REJECTION OF *Ss* HOLDING EACH POSITION

(Columns show latitude of acceptance (strongly accept plus accept), latitude of rejection (strongly reject plus reject) and positions neither acceptable nor unacceptable to *Ss* holding given positions.)

Rating Positions	Own Position								
	A	B	C	D	E	F	G	H	I
A	√√	√	0	×	×	××	××	××	××
B	√	√√	√	0	×	×	×	×	×
C	0	√	√√	√	0	×	×	×	×
D	×	0	√	√√	√	0	×	×	×
E	×	×	0	√	√√	√	0	×	×
F	×	×	×	0	√	√√	√	0	×
G	×	×	×	×	0	√	√√	√	0
H	×	×	×	×	×	0	√	√√	√
I	××	××	××	××	×	×	0	√	√√

Code: √√ = strongly accept; √ = accept; 0 = neither accept nor reject; × = reject; ×× = strongly reject.

Mean frequencies in above table:

	A,B,G,H,I	C,D,E,F
√ and √√	2.7	3.0
0	1.3	2.0
× and ××	5.0	4.0

change as a function of the relative distance between *S*'s stand and the position of communication on an ego-involving issue. Our findings suggest that the relative distance between *Ss*' attitudes and position of communication may be useful in explaining apparently contradictory effects of communication in producing attitude change in the intended direction, no change, and change in the opposite direction. This approach seems to have predictive value for *S*'s reactions to other positions on an issue and susceptibility to change once his own stand is ascertained.

On the basis of results in Table 1 and evidence from previous studies (7, 15), we can formalize the pattern of acceptance and rejection for various positions by *Ss* holding each initial position. In Table 4, the expected response to each opinion item is given for *Ss* checking positions from A to I as their most acceptable stand. The *S*'s own stand is indicated by √√, "also acceptable" with √, completely unacceptable with ××, other stands unacceptable with ×, and not clearly acceptable or unacceptable with 0. Thus, the table presents a hypothetical pattern for the "latitude of acceptance" (consisting of *S*'s own position and other acceptable positions) and "latitude of rejection" (consisting of all unacceptable positions) for *Ss* holding each position. Where sufficient numbers of *Ss* are available for a given stand, the empirical distributions of responses show a close correspondence to these hypothetical patterns.

This model for latitudes of acceptance and rejection can be used to explicate results on reaction to communication. By comparison with Figure 2, showing placement of the position advocated in the moderate communication, it can be seen that "contrast effects" begin to appear when the position advocated falls within *S*'s latitude of rejection, an area far removed from his own stand. If the position advocated in communication falls at the limits or slightly beyond the latitude of acceptance, it would be more likely to be assimilated to the latitude of acceptance, although this trend is less clear in the present study.

If we assume that evaluation of communication on an issue in which *S* is personally involved depends on two factors: (*a*) what the position of communications is perceived to be and (*b*) what segment of the acceptance-rejection range is associated with the perceived position, we can predict from Fig. 2 and Table 4 the degree of favorableness of evaluations of the communication (Fig. 1). When the communication falls within the latitude of acceptance (√ or √√), it is judged fair and unbiased. As the perceived position of communication moves outside the latitude of acceptance into the latitude of rejection (×), it is increasingly considered unfair and propagandistic. For example, an individual checking B as his most acceptable stand displaces the F communication to G ("contrast effect") which, according to Table 4, is clearly in his latitude of rejection; an *S* checking A as his own stand places the F communication between G–H, still further toward the extreme end of the latitude of rejection, and the percentage of favorable evaluations of the communication declines still further (cf. Fig. 1).

Two general problems are suggested by the present analysis. Since *S*'s stands on an issue of personal concern produce variation in reaction to communication and attitude change, the effect of the degree of ego-involvement in issues needs exploration. It is expected that latitudes of acceptance and rejection will vary as a function of the degree of ego-involvement, and these variations may affect the

ranges in which assimilation and contrast effects in judging communication occur. A person strongly committed on an issue will be more discriminating in accepting stands (raised threshold of acceptance) while one who is less committed will be willing to consider a larger number of alternative positions (lowered threshold of acceptance). Corresponding effects on extent of opinion change with different distances between *S* and communication are to be expected. With low involvement issues one would anticipate increase in opinion change with increased separation over a considerable range, whereas with high involvement issues one would expect opinion change over only a narrow range of separation, with resistance to acceptance of the communication for the remaining distances.

Finally, an interesting problem remains as to the role of the communicator in influencing the range of assimilation. In the Hovland-Pritzker study (6) the communicator was an authority respected by *S*s on the issue presented. There the greater the distance between communicator position and that of *S*, the greater the opinion change. In the present study, on the other hand, the communicator was an anonymous individual whose acceptability might be determined by the stand taken by him on the issue under discussion. If he differed in position only slightly from *S* he would be regarded as fair and authoritative and bring about a shift in opinion, but if he differed greatly he would be regarded as incompetent and biased and fail to influence *S*'s opinion. To explore this factor and the one of involvement mentioned in the preceding paragraph, a factorially designed experiment will be required in which type of communicator (positive, negative, and neutral), type of issue (high involvement and low involvement), and distance between the communicator's stand and *S*'s own stand, are all systematically varied to permit analysis not only of the main effects but also of the more interesting interaction effects.

SUMMARY

The relationship between the attitude of the recipient and the position advocated in a communication was studied under conditions where a communicator not known to *S* presents a point of view on a controversial issue which differs from that of *S* by varying amounts.

The topic discussed was prohibition of alcohol. The *S*s came from a dry state where this was a lively issue. Each *S* indicated for a series of nine statements ranging from extreme dry to extreme wet: (*a*) the position most acceptable to him, (*b*) other acceptable positions, (*c*) the position most objectionable to him, and (*d*) other objectionable positions. Groups of *S*s were selected whose stands on the issue were known (e.g., WCTU members) for comparison with unselected groups of college students. In all, 514 *S*s were used. In a subsequent session "wet" groups received a "dry" or "moderately wet" communication; "dry" groups received a "wet" or "moderately wet" communication; unselected groups received a "dry," "wet," or "moderately wet" communication. After the communication, opinion measurements and reactions to communication were obtained. The findings were as follows:

1. When the distance between *S*'s own stand and the position advocated in communication is small, the communication is judged favorably, as fair and factual. With increasing distance, the favorable reaction is sharply reduced and the communication perceived as propagandistic and unfair.

2. The *S*s whose own stands diverge widely from the position advocated perceive the communication as further removed from their own stand than it is ("contrast effect"). The present results indicate, though less clearly, that *S*s whose own stands are close to the position advocated perceive the communication as closer to their own stand than is the case ("assimilation effect").

3. The most frequent result for *S*s whose own stand diverges widely from that advocated in communication is to remain unchanged in their initial attitudes. More *S*s with moderate positions closer to the stand in communication changed in the direction advocated.

It is suggested that the relative distance between *S*'s own attitude and communication along with *S*'s latitudes of acceptance and rejection for various stands on the issue may provide a basis for predicting reactions to communication and susceptibility to change. Further research varying degree of ego-involve-

ment in issues and attitudes toward the communicator is suggested.

REFERENCES

1. FRENCH, J. R. P., JR. A formal theory of social power. *Psychol. Rev.*, 1956, **63,** 181–194.
2. GOLDBERG, S. C. Three situational determinants of conformity to social norms. *J. abnorm. soc. Psychol.* 1954, **49,** 325–329.
3. HEINTZ, R. The effect of remote anchoring points upon judgments of lifted weights. *J. exper. Psychol.*, 1950, **40,** 584–591.
4. HOVLAND, C. I. Effects of mass media of communication. In G. Lindzey (Ed.) Ch. 28, 1062–1103, *Handbook of social psychology.* Cambridge: Addison-Wesley, 1954.
5. HOVLAND, C. I., JANIS, I. L., & KELLEY, H. H. *Communication and persuasion.* New Haven: Yale Univer. Press, 1953.
6. HOVLAND, C. I., & PRITZKER, H. Extent of opinion change as a function of amount of change advocated. *J. abnorm. soc. Psychol.*, 1957, **54,** 257–261.
7. HOVLAND, C. I., & SHERIF, M. Judgmental phenomena and scales of attitude measurement: Item displacement in Thurstone scales. *J. abnorm. soc. Psychol.*, 1952, **47,** 822–833.
8. KNOWER, F. H. Experimental studies of changes in attitudes: I. A study of the effect of oral argument on change of attitude. *J. soc. Psychol.*, 1935, 315–345.
9. LUBIN, A., & FISHER, S. Distance as a determinant of influence in a two-person continuous interaction situation. E.P.A. Talk. Title in *Amer. Psychologist*, 1956, **11,** 491.
10. MANSKE, A. J. The reflection of teachers' attitudes in the attitudes of their pupils. In G. Murphy, L. B. Murphy, & T. M. Newcomb (Eds.), *Experimental social psychology.* New York: Harper, 1937.
11. MURPHY, G., MURPHY, L. B., & NEWCOMB, T. M. *Experimental social psychology.* New York: Harper, 1937.
12. REMMERS, H. H. Propaganda in the schools: Do the effects last? *Publ. Opin. Quart.*, 1938, **2,** 197–210.
13. ROGERS, S. The anchoring of absolute judgments. *Arch. Psychol.*, 1941, No. 261.
14. RUSSELL, D. H., & ROBERTSON, I. V. Influencing attitudes toward minority groups in a junior high school. *School. Rev.*, 1947, **55,** 205–213.
15. SHERIF, M., & HOVLAND, C. I. Judgmental phenomena and scales of attitude measurement: Placement of items with individual choice of number of categories. *J. abnorm. soc. Psychol.*, 1953, **48,** 135–141.
16. SHERIF, M., TAUB, D., & HOVLAND, C. I. Assimilation and contrast effects of anchoring stimuli on judgments. (Mimeo. Rep.)
17. WEISS, W., & FINE, B. J. Opinion change as a function of some intrapersonal attributes of the communicatees. *J. abnorm. soc. Psychol.* 1955, **51,** 246–253.
18. WILKE, W. H. An experimental comparison of the speech, the radio, and the printed page as propaganda devices. *Arch. Psychol.*, 1934, No. 169.
19. WILLIAMS, A. C., & REMMERS, H. H. Persistence of attitudes concerning conservation issues. *J. exper. Educat.*, 1939, **40,** 354–361.
20. WILLIAMS, R. M., JR. *The reduction of intergroup tensions.* New York: Soc. Sci. Res. Council, Bull. 57, 1947.

Received November 30, 1956.

COMPATRIOT AND FOREIGNER: A STUDY OF IMPRESSION FORMATION IN THREE COUNTRIES[1]

JEROME S. BRUNER
Harvard University
AND
HOWARD V. PERLMUTTER
Massachusetts Institute of Technology

A COMMONSENSE observation is the starting point of our inquiry. If one were to enter a room and be introduced to a Frenchman, a German, and an American, it would be likely that one's first impressions of the three would be such as to magnify the nationality difference between our new acquaintances. The Frenchman might appear more "typically French," the German more "typically German," and the American more "typically American." Our perceptual selectivity would register cues that differentiate the three in terms of their most evident difference: their nationality. To this commonsense observation we must add a second. It is also likely that, if one were an American, the Frenchman and German would seem more "typical" of their nationality than would the American to whom we had been introduced. And this would perhaps be the more so to the degree that we were well acquainted with Americans but had experienced little contact with Frenchmen or Germans.

The matter need not be limited to nationality. Introduced at the same time to two men, one as a professor and one as a businessman of our own nationality, likely as not we would seek the cues that permitted maximum differentiation of the two in terms of the difference of profession. And again, the degree to which this would be so might depend upon how well acquainted we were with practitioners of the two professions: the better acquainted, the less the likelihood of role stereotyping. If we were familiar with one of the professions, by the same token, the other would undergo a stereotyping in our impression formation.

These commonsense observations; or rather, the hypotheses to which they give rise and with the testing of which this paper is concerned, lead to several thoughts about the nature of the process that has come to be called "the formation of first impressions." It suggests in the first place that under certain conditions, our impressions have the effect of accentuating the differences between people encountered briefly, and further, that there are other conditions that reduce this tendency to the accentuation of difference. Let us state the matter more formally in terms of a set of hypotheses.

Hypothesis 1. If a person be placed in a particular category on first encounter (nationality, role, occupation), the more widespread one's experience with diverse members of the category the less will this category membership affect the impression formed of the person encountered. Put concretely, if one knows Americans better than Frenchmen or Germans, the first impression formed of an American will be less determined by his nationality than in the other two cases. Or if one knows musicians better than theoretical physicists, the former occupation will affect a first impression less than the latter.

Hypothesis 2. If three pieces of wood, alike in all respects save their grain, be viewed simultaneously, an observer will be more likely to notice grain characteristics in the wood than if one of the pieces is presented alone. Put more formally, that attribute which distinguishes an array of objects will be most salient in viewing the array. Given a single object, there are no similar constraints on what an individual may notice (3). With reference to our own experiment, if the individual is asked to form impressions of an array of individuals differing only in nationality, then nationality will be a more determinative factor in impression forma-

[1] This research was jointly supported by the experimental project of the research program on international communication at the Center for International Studies, Massachusetts Institute of Technology, under the direction of Dr. Perlmutter, and ONR contract ori-0746 at Harvard University under the direction of Dr. Bruner. We are especially grateful to Dr. Renate Mayntz of UNESCO, and M. Michel Hurtig of the Centre National de la Recherche Scientifique of Paris for assistance in the collection and analysis of data.

tion than if the individual is set the task of forming an impression of a single individual.

With this much of an introduction, we may turn to the study designed to examine and test these hypotheses.

DESIGN OF THE STUDY

The basic design of the study was simple in outline. By the use of a brief sketch, a stimulus person or stimulus persons were presented to our subjects (*Ss*). The *Ss* were then given the task of checking on a trait list those items they thought characteristic of the stimulus person as a "first impression." The stimulus persons presented varied in terms of their nationality (French, German, or American) and in terms of their occupation—unspecified, college professor, or businessman. Moreover, the impression formation situation varied, as required by our second hypothesis, in terms of whether the *S* had a single person of a particular nationality in mind when forming his impression or had in mind three people of different nationality: what we shall call the "single-impression" and the "triple-impression" situations. After filling in the impression trait list, *Ss* were asked to indicate, in a manner to be described, the bases upon which they had decided to check various traits. The *Ss* were all university students and themselves varied in nationality. There were university students from Boston; Paris and Dijon, in France; and a set of groups from Berlin, Hanover, and Cologne, in Germany. Students in each country, then, were forming an impression of a compatriot and/or of a foreigner. We may consider now the specifics of the design.

Thirty-seven groups of *Ss* were used in the "single-impression" situation, a third of them Germans, a third French, and a third American. The total sample was comprised of 576 *Ss* of whom 298 were male and 278 female. The groups who were given the test while assembled, varied in size from 13 to 29 *Ss*. Each *S* in these groups, as we have noted, formed an impression of a single stimulus person. The twenty-seven groups and the number of *Ss* in each is set forth in Table 1. Since no consistent sex differences were found, the data for both sexes were combined.

Only nine triple-impression groups were required to fulfill the design. The total population in these nine groups comprised 218 *Ss* of whom 108 were males and 110 females. Each *S* formed impressions of three stimulus persons of different nationality: in a third of the groups the stimulus persons were a French, American, and German professor; in another third, French, American, and German businessmen; and the remaining third, a French, American, and German person of unspecified occupation. The design and the number of *Ss* in each group is presented in Table 2. These *Ss* knew in advance, of course, that they were to form impressions of people of three different nationalities.

TABLE 1

THE NUMBER OF SUBJECTS IN EACH OF TWENTY-SEVEN SINGLE-IMPRESSION GROUPS

Subjects	Stimulus persons who are—								
	American			French			German		
	Prof.	Bus. man	Unspec.	Prof.	Bus. man	Unspec.	Prof.	Bus. man	Unspec.
American	19	13	21	15	18	22	16	18	19
French	29	22	18	29	28	27	28	28	27
German	20	20	20	20	20	19	20	20	20

TABLE 2

THE NUMBER OF SUBJECTS IN EACH OF NINE TRIPLE-IMPRESSION GROUPS

Subjects	Stimulus persons who are—		
	Amer., Fr., & Ger. Profs.	Amer., Fr., & Ger. Bus. men	Amer., Fr., & Ger., Unspec.
American	26	28	30
French	26	25	23
German	20	20	20

All *Ss* were given a booklet with the instructions not to look ahead in the questionnaire and to finish each page before turning to the next. Both single-impression groups and triple-impression groups were told: "You will be asked to form impressions of specific persons in this questionnaire." The latter were further instructed: "These people are from different countries."

The instructions and description of each stimulus person were as follows:

> The object of this test is to determine the extent to which people are capable of judging a person from a few facts about that person. A brief characterization of a specific individual appears below. Read it care-

TABLE 3
THE TRAIT LIST AND ITS GERMAN AND FRENCH EQUIVALENTS

Synonyms for "Energetic"					
Focused			Diffuse		
Trait	German	French	Trait	German	French
resolute	entschieden	résolu	bustling	geschäftig	agité
tenacious	zäh	tenace	animated	angeregt	animé
dogged	verbohrt	persévérant	spirited	munter	fougueux
determined	entschlossen	décidé	vivacious	lebhaft	vif

Synonyms for "Intelligent"					
Focused			Diffuse		
Trait	German	French	Trait	German	French
analytic	analytisch	analytique	wise	weise	sage
brilliant	geistreich	brillant	judicious	abwägend	judicieux
quick	schnell	alerte	level-headed	einsichtig	bien équilibré
logical	logisch	rationnel			
penetrating	durchdringend	perspicace	prudent	klug	prudent
astute	scharfsinnig	astucieux	sensible	vernünftig	sensé
clever	gescheit	adroit	sound	gesund	doué de bon sens
discerning	scharfblickend	doué de discernement	unaffected	ungeziert	simple
canny	schlau	avisé	sensitive	empfindsam	pourvu de flair
shrewd	listig	sagace			
cunning	verschlagen	ruse	tactful	taktvoll	diplomate
calculating	berechnend	calculant d'avance	understanding	verständnisvol	comprehensif
			unprejudiced	vorurteilsfrei	impartial

Synonyms for "Well-adjusted"					
Focused			Diffuse		
Trait	German	French	Trait	German	French
self-possessed	gelassen	maître de lui	cheerful	frölich	joyeux
			happy	glücklich	heureux
philosophical	philosophisch	philosophe	satisfied	zufrieden	content de bon sort
deep	tief	profond			

fully and try to form an image of this person. You will be asked to record your impressions.

He is a very typical _____. There is general agreement among those who know him that he is intelligent, energetic, and well-adjusted. Now 42 years old, he is married and lives in a large city in _____."

For the American businessman, the *S* would be told:

"He is a very typical *American businessman* . . . who lives in a large city in the *United States*." For the French college professor, the subjects would be told: "He is a very typical *French college professor* . . . who lives in a large city in *France*," etc.

Next, *S*s were asked to turn the page and choose from a list of 38 traits (see Table 3, which contains the trait names in English, French, and German) those which "best characterized the person described." The *S*s were allowed to choose as many as they wished.

After the trait list had been checked (in the case of the triple-impression group, after trait lists for all three stimulus persons had been checked), *S*s were next asked the bases on which they had checked specific traits on the list. A series of letters were used to describe these *determinants* of each trait-choice. For each trait checked, *S*s wrote:

E if the trait *energetic* of the stimulus person contributed to that choice,

I if the trait *intelligent* of the stimulus person contributed to that choice,

W if the trait *well-adjusted* of the stimulus person contributed to that choice,

V if the *social role* contributed to that choice,

N if the *nationality* of the stimulus person contributed to that choice.

Subjects could use *any* combination of the letters to describe completely the bases for each choice. They were encouraged to specify freely any other determinants of their choice, although few actually did.

Since we shall rely heavily upon the report of *Ss* concerning the determinants of their choice of traits, a word about the "meaning" of such reports is necessary. One cannot naively assume that *Ss* "know" what led them to check a particular trait. It is as conceivable that the degree of one's hostility toward a businessman-father is as much a determinant of choice as anything contained in the sketch. But it would be just as naive to assume out of hand that *Ss*, university students in this case, are completely incapable of "knowing" the basis for their selection of traits. The issue is not one that can be resolved. Yet, we have asked our *Ss* to indicate the basis of their choices. We take their responses as symptomatic not of the "true" basis of choice—whatever that may mean ontologically—but rather as *a* basis for inferring what underlies their choice. In fact, the only proper basis of inference would be further systematic variation in the trait lists used, in the instructions given *Ss*, and in the nature of the responses they gave by which we make our inferences about causes. We shall go ahead in this paper treating our *Ss*' reports on determinants *as if* they could be taken as a proper basis of inference. In a final section we shall reconsider the matter.

Design of the Trait List

The traits used to characterize the stimulus persons were constructed in the following way, the activity being conducted in English: (*a*) A list of synonyms of the stimulus traits "energetic," "intelligent," and "well-adjusted" were gathered; (*b*) synonyms were divided into dichotomized groupings on the basis of certain dimensions. (See Table 3.) Synonyms of the stimulus trait "energetic," for example, were subdivided into "focused-energetic" and "diffuse-energetic." Under "focused-energetic" we put the traits: bustling, animated, spirited, and vivacious. These dichotomies were based on hunches concerning shifts in the meaning of stimulus traits when connected with persons of various nationalities.

The final ordering of traits in the list was random. Recall that *Ss* were asked to report the determinants of each trait they checked. With the exception of two traits all traits on the list were found to be determined by the stimulus traits for which they were designed to be synonyms. So *Ss* checking "sensible" on the trait list would more often report that the "intelligent" characteristic of the stimulus person determined their choice rather than the fact that he was said to be "well-adjusted" or "energetic."

In both France and Germany the researchers were given the American trait list and were asked to translate this list into their respective languages. A number of problems arose. First, it was difficult to establish an exact lexical equivalent because either the language did not utilize the stimulus traits or the check list traits in the same way, or the lexical equivalent was not used with the same degree of frequency in each country according to the judgment of the respective researchers. In all cases, the list was translated by not less than three bilingual individuals native to the country, and a final discussion was held to decide which translation of the equivalent trait in the American trait list was to be used.

There is no practicable procedure for assuring that the trait lists in the three languages are either denotatively or connotively identical, and the results to be reported must be evaluated in the light of this fact. A close approximation to a test of identity is by the use of a matching procedure. We have asked English-German and English-French bilinguals to match items in the two lists to ascertain whether, within this universe of traits, identity matches could be obtained. But this does not obviate the contingency that the items matched across languages are not the *closest* synonyms possible. All that we can say is that we have taken precautions. For a fuller account of the problems of translating traits from one language into another, the reader is referred to the reports of Perlmutter, Mayntz, and Hurtig, (6) and Lenneberg and Roberts (4).

RESULTS

One of our hypotheses was that *Ss* operating in the triple-impression situation will rely more heavily on nationality as a determinant than

will the single-impression *Ss*. A gross test of the hypothesis is provided by comparing the number of times on the average that *Ss* in the two situations indicated that their choice of a trait for characterizing the person was based upon knowledge of nationality. Grouping together all French, American, and German stimulus persons without regard to occupation and comparing the average number of times that *Ss* justified their choice of a trait by reference to nationality, we obtained the confirmatory results presented in Table 4.

TABLE 4
AVERAGE NUMBER OF TIMES THAT TRAIT IS CHECKED ON BASIS OF NATIONALITY OF STIMULUS PERSON

Group	French Subjects		German Subjects		American Subjects	
	Foreigner	Compatriot	Foreigner	Compatriot	Foreigner	Compatriot
Sing.-Impress.	4.6	2.8	3.9	1.4	3.3	1.3
Trip.-Impress.	6.8	6.8	6.6	4.5	5.2	4.3

Another way of stating these results is by reference to comparisons between specific single- and triple-impression groups. There are 27 single-impression groups as indicated in Table 1. For each, there corresponds a comparable record given under triple-impression conditions. In all 27 of the possible comparisons, the triple-impression group shows a greater reliance on nationality as a determinant of their responses, a result of a confidence level well beyond the .01 level as determined by the Dixon-Mood Sign Test.

Table 4 also confirms our other hypothesis that, in general, one uses nationality determinants less often in forming one's image of a compatriot than in forming an impression of a foreigner. This result was obtained in both single-impression and triple-impression situations. Comparisons may be made in this way. Take first the single-impression situation. For each group of *Ss*, let us compare the use of nationality as a determinant in forming an impression of a compatriot and of a foreigner. Thus, we will compare the use of nationality determinants for Americans judging Americans with their use of this determinant in judging Frenchmen and in judging Germans. We thus have two comparisons. We may also compare the use of nationality when Americans form an image of an American professor as compared with their usage in forming an impression of a German professor and a French professor. For *Ss* of all nationalities, 18 such comparisons are possible for the single-impression situation and another 18 for the triple-impression procedure. In 34 out of the 36 comparisons thus afforded, nationality is used more frequently as a determinant in forming an impression of a foreigner than of a compatriot, again a statistically highly reliable result.

Table 4 indicates, however, that under single-impression conditions, this tendency to use nationality more often on foreigners than on compatriots is considerably more marked than under triple-impression conditions.[2] The latter conditions seem to have the effect of making nationality salient even in sizing up a fellow national. We may note, for example, that with the French *Ss* operating under triple-impression conditions, nationality determinants are used as often on compatriots as on foreigners. It was in this subgroup that the two reversals in trend noted above were found.

We come now to several corollaries of the major hypothesis with which we have just been concerned. The first of these has to do with the specificity of the category into which a stimulus person is "coded." Recall that *Ss* were given sketches of a college professor, a businessman, and a person of unspecified vocation varying, of course, in nationality. Recall also that the *Ss* are themselves university students in working contact with professors perhaps more than with businessmen. It would follow then that the category "professors" would be more differentiated for them than the category "businessmen." We would hypothesize, then, that vocation would more often be used as a determinant in forming

[2] This tendency is statistically a highly reliable one. We test significance in the following way. Compare the difference of nationality determinants used in forming an image of a foreigner and of a compatriot under single- and triple-impression conditions. Eighteen such comparisons are possible. For example, for French *Ss* operating under single-impression conditions, we may compare the difference in number of nationality determinants used in forming an impression of a typical American professor and a typical French professor (3.00 − 2.10 = .90) with the difference obtained for the same stimulus persons judged under triple-impression conditions (6.54 − 5.96 = .58). In 17 of the 18 comparisons possible, the differences were greater for the single-impression groups, a result significant beyond the .01 level.

TABLE 5
AVERAGE NUMBER OF TIMES THAT TRAIT IS CHECKED ON BASIS OF VOCATION OF STIMULUS PERSON

Stimulus Person	Sing.-impress. Groups			Trip.-impress. Groups		
	"Prof."	"Bus. man"	Bus. men Checked More Than Prof.	"Prof."	"Bus. man"	Bus. men Checked More Than Prof.
Amer. subjects:						
Amer.	2.1	6.0	+	4.2	3.9	—
Ger.	3.3	5.0	+	3.5	3.4	—
Fr.	3.3	5.8	+	2.8	3.3	+
Fr. subjects:						
Amer.	2.7	6.3	+	2.5	5.9	+
Ger.	4.5	5.9	+	2.9	4.7	+
Fr.	3.9	6.4	+	3.8	5.7	+
Ger. subjects:						
Amer.	3.1	5.0	+	3.1	4.7	+
Ger.	3.6	6.0	+	4.9	4.6	—
Fr.	3.9	4.6	+	2.6	4.4	+

an impression of a businessman than of a professor, at least amongst our university-based subjects. Such is indeed the case. Consider first the nine single-impression groups. For American *Ss*, for example, we can ask whether vocation is more often used as a basis for forming an impression of an "American businessman" or an "American professor." We can also compare German professor and businessman and French professor and businessman for this group of *Ss* and for *Ss* of other nationalities. This gives us nine possible comparisons. In all nine of these comparisons, the results come out as expected: vocation is more frequently named as a basis of checking list traits for businessman than for professor. These results are summarized in Table 5 where the material for the triple-impression condition is also presented. In the latter case, the picture is less clear: six of the nine comparisons are in the expected direction. If the 18 comparisons be grouped, the resultant figure of 15 in 18 is significant by the Sign test at the .01 level. If the reader examines the reversals in Table 5, he will note that they are more of the order of ties than of reversals in two of the three cases.

THE NATURE OF THE IMPRESSIONS FORMED

Our object in the present report is to consider principally the nature of the impression-forming process and the factors that may influence this process. In the preceding section, the emphasis has been upon determinants, and in the one that follows it will be upon the facets of the images formed by our *Ss* as inferred from the traits they actually checked in characterizing the various stimulus persons presented.

A full presentation of the traits ascribed to Americans, Frenchmen, and Germans of different occupation by respondents of these several nationalities would be a forbidding undertaking. There are 36 groups of *Ss*, each with a somewhat different kind of stimulus person, at least different in terms of conditions under which presented. And there are 38 traits which may be checked.

One approach is to choose a particularly interesting stimulus person—one whose image seems to loom large in the political and social sphere: the American businessman. We will limit our discussion to him as illustrative.[8]

To reduce the complexity of the image that emerges, we limit ourselves to a discussion of those traits that are checked by at least 50 per cent of respondents from a given country. These we may regard as "consensual" impressions. In Table 6 are set forth those traits in which such a consensus was found with respect to the image of the American businessman. All three nationalities, regardless of the method of testing, agree that the American businessman is "determined," and there is also rather wide agreement that he is "sensible." French respondents rather agree with their American fellows in seeing him as "satisfied" and "level headed"—rather a sobersided image—whereas the Germans appear to emphasize a constellation of "resolute," "tenacious," and "calculating." It is difficult, of course, to extract an organized image from a set of checked items on a trait list; and also from the items that go unchecked. For as Asch has so wisely pointed out and demonstrated (2), traits fuse together in an organization, and the organization is often dominated by a single trait. We have no way of determining from our data, for example, whether "calculating" is the organizer of many German impressions, or "resolute" of French images.

There is not a very sharp difference in the images formed by different nationalities, a fair and considerably above-chance agreement

[8] A detailed qualitative analysis of the German findings has been published in Germany (6). In this study, individual differences in the formation of impressions of foreigner and compatriot are discussed.

TABLE 6

TRAITS AGREED UPON BY HALF OR MORE OF SUBJECTS OF EACH NATIONALITY IN CHARACTERIZING AN AMERICAN BUSINESSMAN

Single-Impression			Triple-Impression		
Americans	French	German	Americans	French	German
cheerful	joyeux		cheerful	joyeux	
tactful	diplomate		tactful		
sound	doué de bon sens				
satisfied	content de bon sort		satisfied	content de bon sort	
logical	rationnel				logisch
	résolu	entschieden		résolu	entschieden
	tenace	zäh		tenace	zäh
	simple	ungeziert	unaffected		ungeziert
bustling		geschäftig			geschäftig
sensible	sensé	vernünftig	sensible		
determined	décidé	entschlossen	determined	décidé	entschlossen
clever			clever	adroit	
happy			happy	heureux	
natural					naturlich
level-headed	bien équilibré		level-headed	bien équilibré	
understanding					
	pourvu de flair				
	astucieux				
	maître de lui			maître de lui	gelassen
	avisé			avisé	
	calculant d'avance			calculant d'avance	berechnend
			shrewd		
				alerte	schnell
				animé	
					abwägend
					scharfblickend

being the rule for Americans, Germans, and Frenchmen checking a trait list for stimulus persons of any nationality. We shall not go into detail on this matter, for it is a subject with little issue. Partly it is an artifact of synonymy: an American who is "intelligent" and "well-adjusted" as presented in a sketch is more likely than not to be called "astute" by a fellow-American, "scharfsinnig" by a German, and "astucieux" by a Frenchman. But, then, so is a German stimulus person so characterized.

RECAPITULATION

Our first and perhaps most general hypothesis is that if objects that are alike in all respects save one are considered *together*, their difference in this one respect will be more critical in the impression one forms of the objects. Three identical triplets, differing only in the color of tie they are sporting, will be seen and interpreted more in terms of their tie-wearing habits than would be the case if each one were encountered singly and without the possibility of a simultaneous comparison. A corollary of this hypothesis to which we have addressed ourselves is: If in forming impressions of foreigners and compatriots one is thinking in a comparative context, with the different nationalities in mind while forming one's impression, then the degree to which nationality will influence the impressions formed will be increased. The comparative context, often celebrated rather uncritically as a facilitator of panhuman perceptiveness, may not always have such an effect. In crosscultural contact, an individual may in fact exaggerate the impressions of differentness between foreigner and compatriot. Indeed, anthropologists have often been criticized for their description of the gulf between members of different cultures, particularly by those who, like Allport (1), have tried to emphasize the communalities in the human condition. It may indeed be the special comparative perspective of the anthropologist that produces this emphasis, if indeed it is such.

Our second hypothesis is similarly simple. If a person or object be "placed" or classified in an undifferentiated category with the members of which one has had little experience, the effect will be for the general properties of the undifferentiated category to have a greater

effect on the impression formed of the individual so placed than if the placement had been in a more differentiated category. Category membership, in brief, will have a more telling effect on the impression formed of one of its members in the degree to which the category is differentiated. More differentiated categories have less "stereotyping" effect than less differentiated ones. Thus, the nationality of a person will stereotype one's impression of him to the degree that the nationality is well or poorly known to one. So too with occupation. A college professor will do little stereotyping on the basis of being told that a man he is meeting teaches at a university. Should he know buinessmen less well, the announcement that the man in question earns his living in commerce will have a greater effect on the impression formed. To put the matter more elegantly, a differentiated category is one in which there is a better representation of the likelihood with which specific traits are associated with all or most of its members.

Our two principal hypotheses leave much ground still unturned in the matter of how one forms impressions of compatriots and foreigners, but our findings do serve to sharpen up a few issues. Impression formation depends in massive degree on categorization processes. In essence, we place a person or thing in a category on the basis of a few minimal cues —like a statement of his nationality or occupation—and then proceed to "run off" along the lines of the higher probability attributes associated with people or events included in the category. As noted in a recent work of Bruner, Goodnow, and Austin (3), the attributes that will be emphasized under these circumstances are dictated at least in part by the nature of the discrimination that the impression-former has to make. If one should be asked for the characteristics of man that distinguish him from bears, one set of attributes will become salient. If the task is to distinguish man from all other species, we may have recourse to the old characterization of man as a featherless biped.

We turn at last to the methodological problem introduced earlier. We have asked our *Ss* first to check certain traits that characterize a person presented in a sketch, and then to indicate the basis for doing so—what in the sketch prompted the choice of a given trait. Much of our reasoning has been based on the latter of these data: notably, the reliance on nationality as a basis for checking traits. It is fashionable to be apologetic about utilizing such direct reports as a basis for drawing inferences. We do not feel that the matter can be resolved within the compass of the data available to us. This much is clear. Highly consistent differences have been found, differences that have a congruency as one goes from comparison to comparison. Either our *Ss thought* that their choice of traits was being determined by certain features of the sketches presented to them or they were in some degree so determined. The proper skepticism requires that the matter be left open pending further research—research designed to compare direct reporting with such other indirect methods of response as one may find in the armamentarium of projective tests or psychophysiological procedures. Any one such procedure would be as suspect inferentially as the direct reports of our *Ss*, no more and no less so. Ideally, inferences should be drawn from a very wide sampling of different kinds of responses. Let the reader suspend judgment in the present case, or better still, let him consider whether at this stage of research it is better to commit a Type I or a Type II error.

One final word is in order. Impression formation is not a separate sort of cognitive activity and we are in the debt of Asch (2) for making this clear. It is a phenomenon that requires close analysis in terms of cognitive theory—whether the theory be associationistic, Gestalt, or the type of information-utilization theory that has informed the present enterprise.

REFERENCES

1. ALLPORT, G. W. *Becoming*. New Haven: Yale Univer. Press, 1955.
2. ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.
3. BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
4. LENNEBERG, E. H., & ROBERTS, J. M. The language of experience: a study in methodology. Suppl. *Int. J. Amer. Linguist.*, 1956, **22**, No. 2 (Whole Mem. 13).
5. MAYNTZ, RENATE, & PERLMUTTER, H. V. Der Einfluss von Nationalitätsidentifikation und Berufsdefinition auf die Vorstellungsbildung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 1956, **8**, 2.
6. PERLMUTTER, H. V., MAYNTZ, RENATE, & HURTIG, M. A methodological note on the use of cross-national trait lists for research. Unpublished manuscript, Center for Int. Studies, M.I.T., Doc. -55/36, Nov. 25, 1955.

Received February 6, 1957. Prior publication.

CASE REPORT

THE TREATMENT OF TWO PHOBIC PATIENTS ON THE BASIS OF LEARNING PRINCIPLES[1]

VICTOR MEYER

Institute of Psychiatry, University of London, Maudsley Hospital

THIS paper is presented as an illustration of the possibility of applying experimental method and some of the principles available in psychological theories of learning to the treatment of psychiatric symptoms. As such, it is somewhat similar to an earlier study by Jones (7). It differs from the latter, however, in that it demonstrates the importance of individual differences in personality structure and the manner in which such differences dictate modifications of therapeutic techniques.

The treatments to be described were based mainly on the principle of primary stimulus generalisation (5). According to this principle, a conditioned reaction is not only evoked by the original conditioned stimulus but also, to some degree, by a series of stimuli more or less similar to the original one. The more dissimilar the new stimulus, the less intense is the conditioned reaction.

Many phobic patients display anxiety so intense when presented with certain environmental situations that they are not amenable to any therapeutic approach. However, the generalisation principle suggests that the presentation of an environmental situation which resembles the original one but evokes a reduced amount of anxiety can be exploited therapeutically. In such a situation, one can attempt to substitute an adaptive normal reaction for the original unadaptive anxiety response. A graded continuum of similar situations can then be employed, moving towards the original one, so that the original unadaptive response can be eliminated. This is essentially one of the methods advocated by Jersild and Holmes (6) for the treatment of children's fear, and its theoretical basis has been elaborated by such writers as Guthrie (4) and Wolpe (8). The latter, among others, has demonstrated the usefulness of these techniques in experiments on animals and has applied them successfully in the treatment of a series of psychiatric patients (9).

[1] Thanks are due Dr. D. Hill, Dr. D. A. Pond, and Dr. J. D. Dewsbery, consultants to the Bethlem Royal and Maudsley Hospitals, for their permission, encouragement, and assistance in carrying out the treatment. A debt is also acknowledged to Dr. M. B. Shapiro, Mr. H. Gwynne Jones, and Mr. J. Inglis for offering valuable suggestions.

SUMMARY OF CASE HISTORIES

For the purposes of this paper, a complete case history of each patient is not essential. A brief summary of the psychiatric notes will suffice.

Case 1

The patient, a married woman of 48 years, was admitted to the hospital for investigation of suspected temporal lobe epilepsy. Apart from blackouts, she complained of an excessive fear of going out.

No family history of psychiatric disorder or epilepsy was reported. The birth and developmental milestones seemed normal. She left school at the age of 15 and had a successful work history up to the time of her illness. Her marriage was fairly satisfactory, but there were no children. Before her illness, she was described as quick-tempered but a good mixer, cheerful, hardworking, and conscientious.

Her blackouts started at the age of 28 during pregnancy. Since then, she had suffered 18 blackouts. They all followed physical exertion or emotional upset and were usually preceded by an aura of dull sensation in the chest. She usually lost consciousness, fell limp, and hurt herself on several occasions. On recovery she vomited and felt "horrible."

At the age of 44 she began having peculiar "feelings." These consisted of epigastric sensations, followed by a "horrible" sensation of being about to fall. These episodes occurred frequently when she was about to go out or when she was outside on her own. Occasionally, when she was unable to control these feelings, she tended to panic and then to suffer one of her blackouts.

As a result of these symptoms, she developed a strong fear of going out on her own, of

traveling on public conveyances and of having sexual intercourse for fear of an attack. From the age of 46 she refused to go out unaccompanied. Shortly before the admission to the hospital, she was occasionally unable to go out with friends or even with her husband. She gradually became more and more anxious, depressed (entertaining suicidal thoughts), and fearful. Losing confidence, she resigned from a responsible job and stopped caring for her home. She continued working until admission, but needed to be accompanied all the way to and from work. On admission she was pleasant, sociable, cooperative, rational, and well oriented.

The results of all investigations were negative with respect to temporal lobe epilepsy. Except for some mild abnormalities shown on the air-encephalogram, all the tests contraindicated organic involvement. In view of the possibility that the attacks might be of cardiac origin, she was referred to cardiologist. He could find no abnormalities in the cardiovascular system.

The patient was diagnosed as exhibiting a phobic state with blackouts of unknown nature and was referred to the Psychology Department for the symptomatic treatment of her phobias.

Case 2

A man of 42 years was admitted to the hospital for investigation of blackouts. He also suffered from excessive tension and various fears, mainly related to entering enclosed and crowded spaces.

Two instances of psychiatric disorder were reported in the family history: His mother had a "nervous breakdown," and his elder brother experienced difficulty in going out alone. The patient's early development and childhood seemed normal. He left school at the age of 14 and worked successfully as a precision grinder. He had been happily married and had two children. Before his illness he was described as undemonstrative and timid, but affectionate as a father and husband.

At the age of 22 he had several fainting attacks on parades when serving in the armed forces. Following these, he developed episodes of apprehension which usually occurred in crowded places. He felt tense, flustered, and experienced an urge to run away in a panic. Strange or crowded places and feelings of boredom were associated with the onset of these episodes. The presence of a friend and strong interest tended to prevent their development. Up to the age of 38, he could preserve control and avert panic. After that time, control became more tenuous. On four occasions, episodes of panic terminated in fainting attacks. As a result, he became worried about himself and sharply curtailed his social activities. Shortly before admission to the hospital, he had a feeling of apprehension in a barber's shop, panicked, mounted his bicycle and shortly afterwards collapsed. After this accident he refused to ride the bicycle and would not go out. He became anxious, tense, and depressed.

He was treated by his own physician with phenobarbitone with very little success. On examination for possible temporal lobe epilepsy, the results were negative. When admitted to the hospital, he appeared very tense and anxious, but rational, well oriented, and cooperative. He was reluctant to talk about himself and needed much encouragement.

Like the first case, he was referred to the Psychology Department for symptomatic treatment of a phobic state.

TREATMENT

Case 1

The patient was interviewed twice with the aim of getting detailed information about her phobic symptoms. The rationale of the treatment was then explained and discussed with her. During this time the experimenter (*E*) endeavored to establish good rapport. According to the theory outlined, this is essential since *E* becomes part of the various environmental situations during treatment and, by establishing an effective relationship, he becomes a reassuring stimulus, tending to reduce anxiety.

No exact record of each treatment session was kept and, therefore, some of the description lacks precision. Since the most likely place for the abnormal behavior to occur was a door leading to the "outside" and the "outside" itself, the hospital roof garden, which has walls but no ceiling, was selected as a starting point for treatment. On the first occasion, the subject (*S*) entered the garden with *E*. She only reported mild "thumping in the stomach" throughout this session. Next day she reported

feeling perfectly well and volunteered to enter the garden on her own. When she had gone twice on her own without any signs of the symptoms recurring, the place of treatment was changed to the main hospital garden. She went out several times with *E* on successive days, then was required to go out on her own to meet *E*, and then to seek and find him. Since *S* appeared rather dependent on *E*'s presence, other staff members of the ward were asked to participate in the treatment without *S*'s knowledge of prearrangement by *E*. She went out with various nurses and patients for walks in the garden. Similar expeditions took place during the evening. Not more than two sessions (30 minutes each) were given each day. When asked to go into the garden on her own, she did without any difficulty. From that point, she was encouraged to take frequent walks into the garden during the daytime and the evenings.

The next stage consisted of *E*'s taking walks with the patient outside the hospital grounds. These started in back streets and were gradually extended into the main street; the distance covered increased gradually day by day. She also went out with other people. Eventually, she was able to set out on her own to meet *E* and others outside. She was taken out at least three times when feeling "upset." She took four bus rides accompanied by *E* and frequently volunteered to go shopping with other patients. After two weeks (16 sessions) of this treatment, she was able to make short expeditions on her own without difficulty.

She spent the subsequent 16 days going for walks with *E* and other people and on her own. Short trips on buses were included, and at least three times she went out on her own in the evening. She spent three weekends at home but was instructed to go out on her own only if she felt confident. She managed to take three short walks near her house. On the last day of treatment she traveled some three miles away from the hospital and back. During the whole treatment, she never reported any symptoms except occasional mild thumping in the stomach. Her behavior on the ward improved. According to the psychiatrist in charge and the nursing staff, she had very few "upsets" and was more cheerful. She felt confident and eager to be discharged.

The treatment took about five weeks, and the patient claimed to have enjoyed it. When she was discharged, her relatives and friends were told that she should be encouraged to go out on her own provided she felt confident and not upset. It was planned that she should be seen once every two weeks in the outpatient department for follow-up purposes and, if necessary, to modify management.

Follow-up. For nearly five months the patient worked regularly and was unaccompanied when visiting the hospital. She took occasional solitary walks and was able to do her shopping. Each working day she was met at the station by a friend who traveled with her to the factory; on the return journey she was met by her husband. She managed to get to the station on her own, and on two occasions when the friend failed to meet her, she experienced no difficulties. A month after discharge from the hospital, she developed an acute chest pain for which she was treated by her doctor. Also, while on a bus with some friends, she had one of her blackouts without any warning and was unconscious for about 45 minutes. Despite these difficulties and worries, she continued to work and to go out on her own. At the end of the fourth month, she increased the frequency of her unaccompanied expeditions, and for the first time in two years attended adequately to her housework. She also went to work on her own seven times in succession when her friend was on vacation. She reported occasional feelings of anxiety, but she never panicked and was able to ward off the attacks by thinking about treatment and handling a mascot (a piece of candy) which she kept throughout the treatment. The last time she was seen, she felt cheerful, confident about going out on her own, and practically cured of "unreasonable fears."

Three days after the last interview, the patient died following one of her sudden blackouts when travelling with friends. The cause of her death was attributed to left ventricular failure; postmortem examination demonstrated heart disease.

Case 2

Two extensive interviews with this patient indicated that his anxiety was not associated with any specific stimulus but was more generalised and might depend on a variety of

factors. For this reason, a modified technique was adopted, aimed at enabling *S*, by a process of conditioning, to cope efficiently with his anxiety no matter where it arose. First, a course of systematic desensitisation was given, based on Wolpe's (9) relaxation techniques. This form of treatment failed, since the patient reported that he could not evoke any increased anxiety by thinking about stressful situations. He remained tense and anxious throughout the sessions, but reported no changes in feeling; GSRs similarly indicated no disturbances.

The conditioning of anxiety-relief responses was then attempted by means of a technique similar to that described by Wolpe (9). An inductorium with a maximum inflow of 6 volts was applied to two fingers; the strength of the current was controlled by turning a knob. The patient was told to say aloud "calm yourself" when the shock became unbearable. As soon as the patient said this, *E* switched off the current. Wolpe has reported that many of his patients experienced relief from anxiety in disturbing situations when they used these words associated with the release from shock-induced tension.

Since it appeared that *S* was most likely to manifest his symptoms in a crowded cinema, it was decided to initiate a procedure similar to that used with the first patient. It was planned to visit the local cinema at a time when it is usually relatively empty and then gradually delay visits, going later and later each day until eventually the visit would be made at a time when the cinema is commonly quite full. The effect of boredom was also taken into account since the patient saw at least part of the same picture each day. During the first week, the patient received 5 or 6 electrical conditioning trials immediately before visiting the cinema with *E*. After some thirty to forty minutes, both *S* and *E* returned to the hospital, where 5 or 6 further conditioning trials were given.

After seven days of this treatment, there was no change in *S*'s behavior. During most of the visits to the cinema, he reported feelings of tension and an urge to leave. Following a weekend at home, he said that he still felt depressed, anxious, and tense, and although the "calm yourself" technique helped him slightly, he considered that he was generally unimproved by the treatment.

In every respect, the patient's reactions to the treatment differed from those manifested by Case 1. Whereas Case 1 reacted very quickly to the treatment and a good rapport was easily established, Case 2 remained aloof and detached.

At this stage, Case 2 was reconsidered. A plausible hypothesis was derived from Franks's study (2) concerning the conditionability of extraverts and introverts and from Eysenck's theory of anxiety and hysteria (1). From these ideas, it seemed to follow that Case 1 would condition better and extinguish less rapidly than Case 2. Similarly, both cases should be more neurotic than average, and Case 1 should be more introverted than Case 2. These inferences were supported by the patients' scores on Franks's eyeblink conditionability measure and on the Maudsley Personality Inventory. On the theory[2] that inhibitory drugs decrease conditionability and heighten satiation effects, Case 2 was given 10 milligrams of dexedrine (excitant) for four days. As soon as the effects of the drug became apparent, he was submitted each day to the same treatment procedure as during the previous week.

Throughout this second stage of treatment, the patient felt cheerful and relaxed and enjoyed going to the cinema. At no time did he display any symptoms. On the second day, he was left alone in the cinema. On the last day, he managed to stay alone in a crowded cinema without any difficulty. During the following weekend he went home but took 10 milligrams of dexedrine each day. He traveled alone on buses, rode his bicycle, and went to a cinema with his wife. He reported no feelings of disturbance. On his return from the weekend, the dosage of dexedrine was reduced by 2.5 milligrams a day with an equivalent amount of placebo substituted without the patient's knowledge. The conditioning procedure and the visits to the cinema on his own continued daily after the administration of the drug and the placebo. On the fourth day, he received 10 milligrams of placebo only. During this stage

[2] There is some evidence to support this theoretical position. Unpublished studies at Maudsley Hospital on amobarbital sodium and dexamphetamine sulfate by Franks and Trouton, and on arecoline and methyl atropine by Franks, Laverly, and Trouton have yielded results consistent with the notions applied therapeutically in the present study.

of treatment, he gradually became tense and anxious and reported feeling slightly depressed. After each treatment session, however, he said that he felt better temporarily and at no time did he feel tense and anxious in the cinema.

During the next weekend at home, he was given one placebo tablet for the Saturday and none for the Sunday. He coped with traveling and riding his bicycle, but felt somewhat tense and anxious. He maintained that the cause of these symptoms was the fear that the treatment might fail and that he might have to stay in the hospital. He expressed a strong desire to leave, saying that he had not felt this desire before the treatment started. He also claimed that his present anxiety and tension had a "different quality since he could find a reason for these feelings."

The first day after his return from this weekend, he was given a final test. He had to travel on a bus some three miles away from the hospital, return, and then visit, unaccompanied, another very crowded cinema. He reported that before the bus ride he felt tense and worried, but on his return, he had recovered, regained his confidence, and felt cheerful. He also enjoyed the cinema and reported that he "never felt better."

The rationale of the treatment was explained to him and every step of the treatment (except the substitution of placebo) was discussed with him. After the final test he was discharged.

Follow-up. It was planned to see the patient once every two weeks. He kept his first appointment only, reporting considerable improvement in travelling and in visiting crowded places. He had also gained confidence in himself and in relation to strangers and his superiors. He still occasionally got his "queer feelings" of worry, strain, and fear of fainting, but these were very mild and not very disturbing. The only difficulty he experienced was when riding his bicycle, but he continued to cycle to and from work. Two weeks later he sent a letter, saying that he could not keep the appointment since he was in financial difficulties. Apart from these difficulties, he was managing very well and had mastered his fear of cycling completely. He could cope easily with any uncomfortable feelings he experienced.

Three months later he sent another letter, informing *E* that he was progressing very well. He had taken his children to the coast for a holiday, cycled to and from work without any difficulty, and visited cinemas regularly.

DISCUSSION

This paper attempts to demonstrate the application of some aspects of learning and the personality theory to the symptomatic treatment of psychiatric patients. At present, the experimental findings are not completely consistent, and the theories themselves are in early stages of development. The available findings, however, suggest rational experimental techniques for symptomatic treatment.

It cannot be strongly maintained that a mere stay in the hospital could account for the recovery from the symptoms, since both patients stayed in hospital well over a month prior to treatment, and according to their verbal account and the observation of their behavior in stressful situations, there was no indication of any improvement. In the present state of knowledge, however, it cannot be argued with any degree of assurance that the improvement of these patients' long-standing phobic symptoms was due to the treatment given. Even if there was sufficient evidence to maintain that the treatment was responsible for the improvement, so many relevant factors were uncontrolled that one would not know which aspect of the treatment was relevant and which irrelevant. These relationships can only be established by more extensive and well-controlled research.

Moreover, the usual objections made against symptomatic treatment on a priori grounds lack force. It seems plausible to argue, as Jones (7) has, that "Much evidence points to the fact that neuroticism is largely a constitutional defect for which no effective radical therapy is yet available. The individual of neurotic constitution in certain environmental circumstances develops certain symptoms. The rational therapeutic approach is then to treat the symptoms and to modify the environment so as to avoid their recrudescence or the development of fresh symptoms. The more specific the treatment, the more likely may be its success."

Several implications of theoretical and practical interest emerge from this study: (*a*) Since both cases manifested blackouts of an

unknown nature, particularly Case 1, the phobic symptoms presented were not certainly known to be functional in origin. The treatment, however, was oriented towards a functional analysis of the disorder. (*b*) It would be desirable to know whether, following a course of treatment like that described, the neurotic reactions are eliminated or merely overshadowed by a stronger normal reaction. Although Wolpe (8) provides some evidence that neurotic reactions in cats can be eliminated, the question must remain open on the present evidence. (*c*) Although on clinical grounds Case 1 was regarded as an hysterical personality and Case 2 as a dysthymic, objective psychological tests indicated the reverse. It seems likely, moreover, that the etiology of the phobic states must have been quite different in the two cases. One would like to know how two patients showing such a considerable difference in conditionability came to develop more or less similar patterns of abnormal behavior. (*d*) The results obtained during the course of treatment are consistent with Eysenck's theory (1) concerning the dynamics of anxiety and hysteria. They also indicate, in accordance with Eysenck's view, that for the purpose of treatment, adequate regard must be paid to the importance of certain individual differences as related to the disorders presented. (*e*) The use of drugs may facilitate and speed up successful treatment of this type, and this theory of individual differences may give a rational basis for the selection of drugs.

SUMMARY

The treatment of two phobic patients has been described. Case 1 manifested an excessive fear of going out on her own; Case 2 displayed disabling symptoms in the form of an excessive fear of going into enclosed and crowded places. Both cases had blackouts of unknown nature.

The treatment program for both cases was mainly based on the principle of primary stimulus generalisation. Owing to the more "diffuse" nature of the anxiety displayed by Case 2, an additional simple conditioning technique was employed.

Case 1 responded to the treatment immediately and an improvement in her general behavior was observed. The effects of treatment persisted for nearly five months, when the patient died of left ventricular failure.

Case 2 failed initially to respond to treatment. An attempt was made to account for this failure in terms of Eysenck's theory of anxiety and hysteria. From the theory, inferences were made with respect to the patient's behavior on objective psychological tests; the results were consistent with the inferences. The treatment was modified according to the theory and the patient responded as expected. Four months after discharge from the hospital, the patient seemed to be managing well and to feel much improved.

REFERENCES

1. EYSENCK, H. J. A dynamic theory of anxiety and hysteria. *J. ment. Sci.*, 1955, **101,** 28–51.
2. FRANKS, C. M., & LAVERLEY, S. G. Sodium amytal and eyelid conditioning. *J. ment. Sci.*, 1955, **101,** 654–663.
3. FRANKS, C. M. Conditioning and personality: A study of normal and neurotic subjects. *J. abnorm. soc. Psychol.*, 1956, **52,** 143–150.
4. GUTHRIE, E. R. *The psychology of learning*. New York: Harper, 1935.
5. HULL, C. L. *Principles of behavior*. New York: Appleton, 1943.
6. JERSILD, A. T., & HOLMES, F. B. Methods of overcoming children's fears. *J. Psychol.*, 1935, **1,** 75–104.
7. JONES, H. G. The application of conditioning and learning techniques to the treatment of a psychiatric patient. *J. abnorm. soc. Psychol.*, 1956, **52,** 414–419.
8. WOLPE, J. Experimental neurosis as learned behavior. *Brit. J. Psychol.*, 1952, **43,** 243–268.
9. WOLPE, J. Reciprocal inhibition as the main basis of psychotherapeutic effects. *Arch. Neurol. Psychiat.*, 1954, **72,** 205–226.

Received October 26, 1956.

CRITIQUE AND NOTES

PERCEPTION OF DISTURBING AND NEUTRAL WORDS THROUGH THE AUTOKINETIC WORD TECHNIQUE[1, 2]

SARNOFF A. MEDNICK, ALAN HARWOOD[3], AND JACK WERTHEIM

Harvard University

THE Autokinetic Word Technique (AWT) presents a situation in which the subject (*S*) is placed in a completely darkened room and asked to read and report words and sentences being written by an apparently moving but actually motionless point of light. A previous study (1) indicates that *S*s do indeed report the perception of words and sentences. The use of the term "perception" is perhaps debatable, but the authors feel justified in its use by the similarity of present operations to those involved in the study of such illusions as the Necker cube.

One of the important uses to which the AWT may appropriately be put is the investigation of nonautochthonous perceptual variables. In this study, it was hypothesized that in the AWT situation "disturbing" words would be perceived with greater latency and less frequency than "neutral" words.

METHOD

A word association test (WAT) of 40 words was administered to 12 male and 12 female paid volunteer college students. All of the words used were drawn from the recent Minnesota standardization of the Kent-Rosanoff Word Association Test on a college population (2). They all have a Thorndike-Lorge AA rating (3), indicating that they are commonly used in the English language. On the basis of probability of response (2), five disturbing and five neutral words were picked for each *S* from his WAT results. In cases where, due to ties, a decision could not be made solely on the basis of probability, the latency of the response served as the deciding factor. The median probability and latency of responses to the disturbing words were .001 and 2.10 sec., respectively; the median probability and latency of responses to the neutral words were .622 and 1.30 sec., respectively.

The *S*s were then introduced to the AWT situation and instructed that we were going to write sentences with the point of light. (These *S*s had had two hours of previous experience in the AWT situation in another study.) Before each sentence was written they were informed that they would be given one of the words in the forthcoming sentence. This word might appear at the beginning, middle, or end of the sentence. The *S* was to inform *E* when he saw this word.

Ten trials were then given in which the stationary pinpoint of light was turned on for 90 seconds along with a cheap electric motor that produced a convincing, variable noise. On five of the trials, neutral words from the WAT were the suggested words. Disturbing words from the WAT were the suggested words on the other five trials. Whenever *S* reported seeing the suggested word during the 90-second period, his response latency was noted. When the light was turned off, *S* was asked to report the complete sentence that had apparently been written by the point of light. *E* was present in the blacked-out room with *S*, who was seated approximately nine feet from the point of light.

RESULTS

The two major measures of this study were (*a*) whether or not *S* reported perceiving a suggested word, (*b*) the latency of his reported perception. Of the 10 suggested words, *S*s reported failing to see a mean of 3.73 words. *S*s failed to see a mean of 2.09 of the five disturbing words and a mean of 1.64 of the five neutral words. This difference, evaluated by the paired replicates test (4), was significant at the .01 level. Thus, *S* failed to report seeing more of the disturbing than neutral words.

The mean latency of report of perception of the suggested words was 58.6 seconds. The perceived disturbing words had a mean latency of 64.4 seconds. The perceived neutral words had a mean latency of 52.7 seconds. The difference between these values was significant at the .01 level by the paired replicates test. Postponement of the report of perception of the disturbing word could be the result either of a simple unfilled delay of response or of the perception of many words preceding the disturbing word. The mean number of words preceding the suggested word was 1.02 words. The disturbing words were preceded by a mean of 1.08 words; the neutral words were preceded by a mean of .96 words, a difference that is not significant. The delay of report of perception of disturbing words cannot therefore be attributed to the perception of other intervening words.

The mean number of words reported in the sentences built around the disturbing and neutral words were 2.54 and 2.64, respectively. This difference, evaluated by the paired replicates test, was not significant.

[1] This research was supported by the Laboratory of Social Relations, Harvard University, and the Social Science Research Council.

[2] Presented to the 1956 meeting of the American Psychological Association.

[3] Mr. Harwood was the recipient of an SSRC Undergraduate Research Stipend, 1956, for work on this study.

SUMMARY AND CONCLUSIONS

Disturbing and neutral words were selected individually for *Ss* by means of a word association test. It was then suggested to these *Ss* that an apparently moving but actually motionless point of light in a completely darkened room (autokinetic effect) would write a sentence containing a given word. The given word was either a disturbing or neutral word.

The *Ss* reported perceiving fewer of the disturbing words. Those disturbing words that were perceived had a greater latency than perceived neutral words.

REFERENCES

1. RECHTSCHAFFEN, A., & MEDNICK, S. A. The autokinetic word technique. *J. abnorm. soc. Psychol.*, 1955, **51**, 346.
2. RUSSELL, W. A., & JENKINS, J. J. *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test.* Office of Naval Res. Tech. Report No. 11, 1954, Contract No. N8.
3. THORNDIKE, E. L., & LORGE, I. *The teachers word book of 30,000 words.* New York: Teachers College, 1944.
4. WILCOXON, F. *Some rapid approximate statistical procedures.* New York: Amer. Cyanamid Co., 1949.

Received October 1, 1956.

LEARNING THEORY AND "OPPOSITE SPEECH"

ARTHUR W. STAATS

Arizona State College at Tempe

LAFFAL, Lenkoski, and Ameen (4), have recently reported the case of a schizophrenic patient demonstrating what they call an "opposite speech" syndrome. The authors state that the syndrome is basically a reversal in certain language usages, especially in the use of "yes" and "no." Feeling that the speech reversal is an expression of repressed hostility, they suggest that opposite speech copes with the hostile impulses by disrupting communication and rejecting other people or by allowing the "verbalization of ideas which the patient consciously rejects" (4, p. 412).

Laffal *et al.* conclude that learning theory has little to offer in this instance or others of pathological speech. ". . . [T]he so-called pragmatic and interpersonal functions of language behavior have been largely neglected by learning theorists in favor of the meaning or semantic aspects" (4, p. 412).

However, a number of recent experimental and theoretical studies (such as 1, 2, 3, 5, 6, 7, 9, 10) have been very much concerned with similar problems of language behavior. Several points can thus be made from a learning theory approach which have relevance for "opposite speech." Skinner, for example, generally defines verbal behavior as "behavior which is reinforced through the mediation of another organism" (7, p. 20). Instead of asking, as do Laffal *et al.*, "what uses [needs] the opposite speech may serve" (4, p. 412), reinforcement theory is concerned with the conditions which increase the probability that the response will occur again in similar stimulus situations or which strengthen the response.

It has been shown experimentally (1, 3, 9, 10) that social reinforcement which follows verbal responses has the effect of increasing their frequency. These studies have used various social reinforcements. The question is thus raised regarding the reinforcements involved in opposite speech. One possibility may be taken from a discussion of generalized reinforcers by Skinner.

> The child who misbehaves "just to get attention" is familiar. The attention of people is reinforcing because it is a necessary condition for other reinforcements from them. . . . The attention of someone who is particularly likely to supply reinforcement—a parent, a teacher, or a loved one—is an especially good generalized reinforcer and sets up especially strong attention-getting behavior (8, p. 78).

It is suggested that attention, operating as a generalized reinforcer, could account in part for the origin and maintenance of opposite speech. The interview quoted by Laffal (4, p. 411) is a good demonstration of how the patient's opposite speech could elicit attention from others. Other reinforcers would also be effective in strengthening reversed language. For example, the patient's reversals were said to "not extend to . . . thought or action. Thus, if asked if he wishes a cigaret, the patient may say, 'No' instead of 'Yes, I do,' but he accepts and smokes the cigaret" (4, p. 412). Giving the cigaret to the "No" response is a reinforcement and would be expected to strengthen that response, i.e., raise the probability that the patient would again say "No" in the same situation. Reinforcements such as these, without con-

comitant admonition, would be unlikely to occur in the normal person's social environment.

Certain implications are derivable from this interpretation. If the opposite speech is maintained by positive reinforcement, then lack of such reinforcement should lead to extinction of such behavior. For example, withholding the cigaret should weaken the strength of opposite speech, and giving the cigaret to correct speech should strengthen that type of response.

Another example of learning theory formulations applicable to opposite speech is offered by Dollard and Miller (2). They discuss repression in terms of reinforcement theory. Certain thoughts arouse anxiety. Cessation of thinking those thoughts reduces anxiety. Thus, stopping thinking about that topic becomes a well-learned response. The same analysis can be applied to speech. Dollard and Miller (2) give an example of a group of people who change the topic of a conversation because it arouses anxiety, and state, "people tend to learn to avoid unpleasant topics of conversation" (2, p. 199).

It could be said that the schizophrenic patient's verbal behavior, when it is not confused by reversal, elicits an anxiety response in him, perhaps because of its typical content. Confused ways of speaking, and perhaps even of thinking, would therefore be anxiety reducing. This rationale could be extended to obsessive thinking in addition to other types of confused schizophrenic speech and thought. In addition, for the schizophrenic, communication with others which is understandable probably introduces touchy subjects which arouse anxiety. Reversed verbal behavior and other confused speech may reduce anxiety when it produces breakdown of the communication and cessation of the anxiety producing subject matter or of the bothersome conversation itself.

Why the schizophrenic's speech might elicit anxiety in him need not herein be elaborated, since the speech symptom is the relevant topic. It could be stated, however, that the unhappy life situation of an adult schizophrenic probably elicits thought and speech which are not positive secondary reinforcers, but instead arouse anxiety. It is also probable that the lack of success of the schizophrenic's life behavior evokes verbal behavior from others which is anxiety producing for the schizophrenic. Verbal and nonverbal behavior of the schizophrenic which would avoid this anxiety would thus be well learned.

In concluding, it should be stated that this note is not intended as a complete analysis of the opposite speech of the schizophrenic patient. Perhaps it points out that hypotheses which apply to pathological language can be derived from a learning theory approach. At any rate, it is suggested that learning theory has reached a state where it has something to offer clinical theory *and practice*. The complex activities to which contemporary learning theory addresses itself, and not without success, indicate that the approach can no longer be ignored.

REFERENCES

1. COHEN, D. J., KALISH, H. I., THURSTON, J. R., & COHEN, E. Experimental manipulation of verbal behavior. *J. exp. Psychol.*, 1954, **47**, 106–110.
2. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
3. GREENSPOON, J. The effect of verbal and nonverbal stimuli on the frequency of members of two verbal response classes. Unpublished doctor's dissertation, Univer. of Indiana, 1950.
4. LAFFEL, J., LENKOSKI, L. D., & AMEEN, L. "Opposite speech" in a schizophrenic patient. *J. abnorm. soc. Psychol.*, 1956, **52**, 409–413.
5. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
6. MOWRER, O. H. The psychologist looks at language. *Amer. Psychologist*, 1954, **9**, 660–694.
7. SKINNER, B. F. *Verbal behavior*. William James Lectures, Harvard Univer., 1947 (mimeo.).
8. SKINNER, B. F. *Science and human behavior*. New York: MacMillan, 1953.
9. TALBOT, E. The effect of note taking upon verbal responses and its implications for the interview. Unpublished doctor's dissertation, Univer. of California at Los Angeles, 1954.
10. VERPLANK, W. S. The control of the content of conversation: reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **51**, 668–676.

Received October 22, 1956.

THE INFLUENCE OF ETHNIC ATTITUDES ON REASONING ABOUT ETHNIC GROUPS[1]

NORMAN M. PRENTICE[2]
Harvard University

IN THE abundant literature on ethnic prejudice, studies exploring the effects of tolerant and intolerant attitudes on reasoning about ethnic groups are rare. In one of the few relevant investigations, Thistlethwaite (5) demonstrated that prejudiced subjects biased their reasoning against ethnic groups. Moreover, an account by Allport (2) of unpublished research by S. B. Sells suggested that tolerant as well as intolerant subjects may bias their reasoning about ethnic groups in directions resonant with their attitudes, a finding in opposition to the traditional characterization of tolerant persons as rational and objective in their appraisal of ethnic groups. The account was nonetheless consistent with the evidence from related attitude areas that opposing social attitudes conduce to distorted reasoning in directions congruent with these attitudes (4). The aim of the present experiment was to examine more rigorously the influence of ethnic attitudes on reasoning logically about ethnic groups. The prediction was made that *tolerant individuals would show bias in favor of ethnic groups while intolerant individuals would show bias against ethnic groups on a test of syllogistic reasoning involving ethnic content.*

METHOD

Test Measures

1. Attitudes. Ethnic attitudes were measured by an 18-item form of the California Ethnocentrism (E) Scale, composed of an equal number of statements disparaging Jews and Negroes.[3] Authoritarian attitudes were appraised by a 12-item form of the California Fascism (F) Scale.[4]

2. Objective reasoning about ethnic groups. The ability to reason logically about ethnic groups was provided by performance on a syllogism test constructed by the author called, "The Woodbridge Reasoning Test."[5] This test consisted of 48 syllogisms randomly arranged which the subjects (*Ss*) rated for logical validity. The *Ss* were instructed to rate the conclusions to the syllogisms as following logically or not necessarily following logically from the premises. Of the 48 syllogisms, 24 were logically valid and 24 logically invalid by this definition. Each group of 24 contained 8 subgroups of three syllogisms. The three syllogisms in each of these groups were alike in validity, syllogistic mood, approximate sentence length, and linguistic style, but unlike in content. The premises and conclusions of one syllogism were tolerant, those of a second were intolerant, while those of a third were neutral; that is, designed not to produce any special affective reaction. The content of none of the affective syllogisms was formally identical to that in the E scale, though both measures dealt with Jews and Negroes exclusively. In contrast with the syllogisms, all of the E scale items were negative; i.e., critical of Jews and Negroes. Content for some affective syllogisms was adapted from Thistlethwaite (5), while that for others was drawn from stereotypes of minority groups enumerated by Allport (2).[6] The following triad in which all the conclusions are logically invalid illustrates the construction:

Tolerant

If no one who is especially loud, noisy, and overbearing is admitted to a country club;
And if many Jews are admitted to a country club;
Then: No Jews are especially loud, noisy, and overbearing.

Intolerant

If no one who mixes especially well with other people is aloof from the community activities;
And if many Jews are aloof from the community activities;
Then: No Jews mix especially well with other people.

[1] This report is based on portions of a doctoral dissertation completed at Harvard University in the Department of Social Relations in 1956. The author wishes to express particular indebtedness to his principal advisers: Drs. G. W. Allport, G. Lindzey, and R. W. White. Drs. N. Kogan and D. J. Levinson generously read this report and offered many valuable suggestions.

[2] Now at Judge Baker Guidance Center.

[3] Items selected from Adorno *et al.* (1) were numbers: 8, 22 (pp. 68–69); 13 (p. 70); 2, 31 (pp. 110–111); 37 (p. 117); and 1–6 of Subscales A and B (p. 142).

[4] The form given was taken from the FERPT Battery used by Dr. D. J. Levinson at Harvard University, Department of Social Relations, 1952. With one exception ("More than anything else, it is good hard work that makes life worthwhile"), the items are either identical or slightly rephrased from Adorno *et al.* (1): b-44 (p. 248); b-21, b-42, c-13, c-25, c-34, c-37 (p. 255); a-9, e-4, f-26 (p. 256); h-18 (p. 257).

[5] This test, together with a table indicating syllogism construction, has been deposited with the American Documentation Institute. Order Document No. 5297 from ADI Auxiliary Publication Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $1.25 for 35 mm. microfilm or $1.25 for 6 by 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

[6] Dr. S. B. Sells graciously offered the syllogism test devised by him previously. Although none of his syllogisms were used in their original form, they provided structural models for the test devised herein. Dr. M. H. Thompson, now of the Philosophy Department, University of Chicago, kindly assessed the correctness of the syllogisms in their final form.

Neutral

If no one who is especially interested in Egyptian art is concerned with Mexican art;
And if many dealers are concerned with Mexican art;
Then: No dealers are especially interested in Egyptian art.

Approximately 35 minutes were allowed for completion of the syllogism test.

An error score formula was derived on the assumption that the influence of ethnic attitudes would lead to reasoning errors in predictable directions on syllogisms with particular ethnic content. Thus, tolerant individuals were expected to have more errors on reasoning items that were valid but contained intolerant content and on reasoning items that were invalid but contained tolerant content. These same individuals were expected to have fewer errors on reasoning items that were valid but contained tolerant content and on reasoning items that were invalid but contained intolerant content. The converse was predicted for intolerant individuals. Thus errors in judgment of validity were expected to be systematically rather than randomly distributed across the reasoning items with affective content. The error score used for the analysis was as follows:

$$\text{Error Score} = [(AV - PV) + (PI - AI)]$$

where

AV = Anti-Valid: number of errors on valid syllogisms with intolerant content.
PV = Pro-Valid: number of errors on valid syllogisms with tolerant content.
PI = Pro-Invalid: number of errors on invalid syllogisms with tolerant content.
AI = Anti-Invalid: number of errors on invalid syllogisms with intolerant content.

The use of a difference score in the formula controlled for the level of reasoning ability. Neutral items were included in the syllogism test to provide variety of content and for other purposes, but scores on these items were not used in the analysis herein.

Subjects

The tests were taken anonymously by 128 white Christian male and female freshmen enrolled at the Boston University School of General Education in the early Fall of 1955 who volunteered for a "study of social attitudes."[7] The Woodbridge Reasoning Test, E and F Scales, were administered in that order. None of the *Ss* had had courses in psychology or formal logic.

RESULTS

Product-moment correlations between the E and F scores and the mean error scores on the Woodridge Reasoning Test were −.22 and −.30.

[7] The cooperation of Dr. G. N. Eddy and Mr. R. Smart, Social Science Department, Boston University College of General Education, in making *Ss* available for this study and prior pilot studies, is gratefully acknowledged.

TABLE 1
COMPARISONS OF MEAN ERROR SCORES WITH HYPOTHETICAL ZERO MEAN SCORES ON WOODBRIDGE REASONING TEST FOR INDIVIDUALS HIGH AND LOW ON E AND F SCALES

	N	$\bar{X}$	*t*	*p*
Low E	31	1.74	4.25	<.001
High E	32	.63	1.27	Reversal
Low F	30	1.57	3.75	<.001
High F	29	−.41	−.88	<.19

Note.—Values are for one-tailed *t* test following Lord as cited by Mosteller and Bush (3).

These values are significant beyond the .02 and .01 levels (one-tailed test), respectively. The correlation coefficients are negative since stronger intolerant attitudes (high E and F scores) are associated with less reasoning bias in the tolerant direction.

A subsequent analysis compared the mean error scores on the Woodbridge Reasoning Test made by individuals in the upper and lower quarters on the E and F scales, with hypothetical mean error scores of zero (unbiased reasoning). As the error score formula is applied, positive error means will be associated with the lower quarters of the E and F scales and negative error means with the upper quarters if respondents are biasing their reasoning in a direction congruent with their attitudes. Table 1 presents the results from this analysis. As indicated, in both the lower E and F quarters, significant bias operating *in favor* of ethnic groups was demonstrated (<.001 in both instances). However, in the high E and F quarters, significant bias operating *against* ethnic groups was absent, although the upper quarter F mean tended in the predicted direction. The absence of intolerant bias can be presumptively related to the composition of the sample. While the neutral item mean for both the attitude scales is 4.00, the item mean for the E scale was 2.72. Thus, as a whole, the sample with respect to the E scale was composed predominantly of tolerant individuals. The sample with respect to the F scale was somewhat more normally distributed with an item mean of 3.80. In this connection, an important distinction between relative and absolute position along these attitude continua should be made. In the present study, for example, the upper quarter E group, although extremely intolerant relatively, tended to be slightly tolerant in terms of the total possible range of ethnic attitudes measured by the scale. Indeed, as indicated in Table 1, the High E group showed a tendency to bias in the tolerant direction. Clearly, in considering extremes, it is important to be concerned not only with the samples at hand but with the theoretical universe as well.

DISCUSSION

In the present study, tolerant individuals biased their reasoning about ethnic groups, while Thistlethwaite (5) using a similar method in 1948 reported intolerant bias alone. The present experiment conducted in 1955 in a social climate somewhat altered during the intervening years, used a completely Northern sample lacking highly prejudiced *Ss*. Thistlethwaite used Southern samples which presumably contained many highly prejudiced *Ss*. He compared the direction of reasoning bias in these with Northern samples which may have contained both intolerant and tolerant respondents. As a consequence, the tolerant bias he predicted for his Northern samples may have been attenuated. However, as Thistlethwaite selected his comparison groups on the basis of region rather than on standard attitude scales, more exact comparisons of these related studies cannot be made. Nevertheless, the findings of both studies taken together support the general assumption that individuals with opposing social attitudes distort their reasoning in directions congruent with these attitudes. Moreover, the findings from the present study challenge the conventional emphasis on veridicality in tolerant individuals. Apparently, tolerance is not always associated with superior reality oriented reasoning.

SUMMARY

This study was undertaken to investigate further the relationships between ethnic attitudes and objective reasoning about ethnic groups. It was argued that tolerant individuals would bias their reasoning in favor while intolerant individuals would bias their reasoning against ethnic groups. The findings revealed that tolerant individuals biased their reasoning significantly *in favor* of ethnic groups. However, intolerant individuals did not bias their reasoning significantly *against* ethnic groups. The lack of bias against ethnic groups was presumptively related to the absence of strongly prejudiced respondents in the sample. Biased reasoning on the part of unprejudiced individuals challenged the traditional assumption that tolerance is invariably accompanied by superior reality directed thinking.

REFERENCES

1. ADORNO, R. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. ALLPORT, G. W. *The nature of prejudice*. Cambridge, Mass.: Addison-Wesley, 1954.
3. MOSTELLER, F., & BUSH, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. I. *Theory and method*. Cambridge, Mass.: Addison-Wesley, 1954. Pp. 289–334.
4. SELL, D. E. *The evaluation aptitude test*. Chicago: Psychometric Affiliates, 1952.
5. THISTLETHWAITE, D. Attitude and structure as factors in the distortion of reasoning. *J. abnorm. soc. Psychol.*, 1950, **45**, 442–458.

Received October 26, 1956.

ERRATA

In the article "Perceptual Organization and Intelligence: A Further Study," by Clinton De Soto and H. Leibowitz, which appeared in the November, 1956 issue of this journal (*J. abnorm. soc. Psychol.*, 1956, **53,** 334–337), some incorrectly calculated probabilities of Type II errors were given. Recalculated probabilities are given in Table 1. These are probabilities of Type II errors for 2-tail *t* tests of the difference between mean vocabulary scores of high and low perceivers in case of certain hypothetical true differences in various experimental conditions and groupings, as explained in the original article. The probabilities in Table 1 are substantially greater than those given earlier, but low enough that a complete absence of significant *t*'s would be surprising if a large true difference existed. All *t*'s obtained were nonsignificant, and the average difference was in fact in the wrong direction.

It may be noted that 1-tail *t* tests might have been used, since the direction of the difference was hypothesized. One-tail *t* tests would have had lower probabilities of Type II errors than 2-tail *t* tests for the hypothetical true differences listed, and would also have yielded uniformly nonsignificant results.

TABLE 1

Condition	Difference	Probability of Type II Error
0.01 sec.	3.76	.65
0.01 sec.	5.20	.42
0.10 sec.	3.99	.56
0.10 sec.	5.53	.30
1.00 sec.	3.39	.57
1.00 sec.	4.70	.32
Above combined	1.87	.63
Above combined	2.53	.40
30 sec.	1.87	.72
30 sec.	2.53	.53

In the article by S. J. Korchin and S. Levine, "Anxiety and Verbal Learning," this Journal, 1957, **54,** 234–240, it was erroneously stated that Dr. Korchin is now at the Research Division, Columbus State Institute of Psychiatry, Ohio State University. This statement applies to Dr. Levine. Dr. Korchin remains with the Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital.

ARE THERE GAPS IN YOUR FILES OF APA JOURNALS

THEN HEAR THIS . . .

The American Psychological Association announces a sale during the period October 1957 through March 1958. Of the following journals, all *available* issues in the volumes for the years preceding 1951 will be offered at a price of only 50¢ (foreign, 60¢) per issue:

American Psychologist
Journal of Abnormal & Social Psychology
Journal of Applied Psychology
Journal of Comparative & Physiological Psychology (1947–1950 only)
Journal of Consulting Psychology
Journal of Experimental Psychology
Psychological Abstracts
Psychological Bulletin
Psychological Index (a few complete volumes, some shopworn)
Psychological Monographs
Psychological Review

Not *all* issues in *all* volumes are available. But—ORDER NOW before more back issues go out of print. From our available stock we will complete as much of your order as possible at this reduced price and for this limited period.

Delivery: 6 to 8 weeks • No dealer or quantity discounts

After this sale, for the years preceding 1948, journals will be available only on microfilm and microcard.

Order from:

American Psychological Association
Department BB, 1333 Sixteenth Street, N.W.
Washington 6, D. C.

SALE OF BACK ISSUES

Neurology and Psychiatry

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological survey in addition to general clinical neurology and experimental neurophysiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition the following special branches are dealt with fully: psychoanalysis, individual and analytical psychology; sexology, criminology, alcoholism, and drug addiction as related to psychiatry; psychosomatic medicine; social and industrial psychology and psychiatry, vocational guidance and mental testing; psychometrics; personality testing and the Rorschach test; heredity and statistical studies as they apply to psychiatry. The subjects of mental defect and epilepsy are fully reported.

Annual volume of 1000 pp., including an index classified by author and subject

Subscription price $22.50 per annual volume

The EXCERPTA MEDICA FOUNDATION

The New York Academy of Medicine Building

2 East 103 Street **New York 29, N. Y.**

NOW AVAILABLE

Psychology and Mental Health

A report of the Institute on Education and Training for Psychological Contributions to Mental Health, held at Stanford University in August, 1955.

Edited by CHARLES R. STROTHER

Topics discussed at the Institute:

Training Needs of Psychologists in Community Mental Health

Specialization in Training

Practicum Training

Training for Therapy

Training for Research in the Mental Health Field

Problems of Departmental Organization

Price: $1.75

Order from:

American Psychological Association, Inc.
1333 Sixteenth Street, N. W.
Washington 6, D. C.